40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

BLADEN LIBRARY SCARBOROUGH
31761 02908 1742

Epidemiological and Biostatistical Aspects of Alcohol Use, Alcoholism and Their Complications
OLE-JØRGEN SKOG

Dependence: Uncertainties, Mysteries and Doubts
GRIFFITH EDWARDS

Evolving Knowledge on Genetic Factors in Alcoholism
C. ROBERT CLONINGER & ERIC J. DEVOR

Reinforcement and Drug Self-Administration: Brain Mechanisms Underlying Dependence
ROY A. WISE

The Evolution of Treatment for Alcohol Problems since 1945
WILLIAM R. MILLER

The Law in Addictions: Principles, Practicalities and Prospects
PATRICIA G. ERICKSON

Forty Years of Basic Research: Investment in the Future
HAROLD KALANT

EDITED BY
PATRICIA G. ERICKSON
HAROLD KALANT

ARF Addiction Research Foundation
Fondation de la recherche sur la toxicomanie

Digitized by the Internet Archive
in 2017 with funding from
University of Toronto

https://archive.org/details/windowsonscience00eric

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

188

EDITED BY

PATRICIA G. ERICKSON

HAROLD KALANT

VINCENT W. BLADEN
LIBRARY

SCARBOROUGH COLLEGE

Copyright © 1992 Addiction Research Foundation
All rights reserved
ISBN 0-88868-196-8
Published by Addiction Research Foundation
33 Russell Street, Toronto, Canada M5S 2S1
Printed in Canada

TABLE OF CONTENTS

Foreword i
ROBIN ROOM

Acknowledgements iv

Epidemiological and Biostatistical Aspects of Alcohol Use, Alcoholism, and Their Complications 1
OLE JØRGEN SKOG

Dependence: Uncertainties, Mysteries and Doubts 37
GRIFFITH EDWARDS

Evolving Knowledge on Genetic Factors in Alcoholism 67
C ROBERT CLONINGER & ERIC J DEVOR

Reinforcement and Drug Self-Administration: Brain Mechanisms Underlying Dependence 83
ROY A WISE

The Evolution of Treatment for Alcohol Problems since 1945 107
WILLIAM R MILLER

The Law in Addictions: Principles, Practicalities and Prospects 125
PATRICIA G ERICKSON

Forty Years of Basic Research: Investment in the Future 161
HAROLD KALANT

FOREWORD

IT IS an honor and pleasure to write a foreword to the work of such a distinguished series of lecturers. In their reports from the front lines of their particular facets of science, the first six speakers exemplify the breadth and diversity of addiction research, while the seventh offers a persuasive and well-documented argument for the long-term utility of basic research in the field.

THE OCCASION OF a 40th anniversary adds a historical dimension to the series, and several contributors have considered the historical dimensions of their topics explicitly. The discussions of historical developments of the various fields put me in mind of Skog's discussion in his lecture of the varied shapes of risk functions: linear, convex and U-shaped. Depending in part on the time frame adopted, we can see the development over time of science in the addictions field in terms of a diversity of shapes. If we focus only on the current moment in science, and on the patterns of citation that back it up, we will tend to see a convex curve: the past leading in an inexorable crescendo to the present. But if we take a longer time frame, and consider the past more in its own terms, the curve sometimes looks like a U. This may even be true of the level of scientific output in a field — despite the convex curve of scientific output in general. Thus the curve of scientific output on alcohol in the 20th century, for instance, would look like a lopsided U, bottoming out in the 1930s and 1940s. The U-shaped curve may be more obvious when we turn from the scientific output to the scientific conclusions. In the present series, Edwards discusses this most explicitly: he notes that alcohol withdrawal almost disappeared as an entity in the medical literature in the 1940s. Cloninger and Devor also obliquely refer to a U-shaped pattern: the genetic inheritance of alcoholism was considered well established around 1900; in the 1940s and 1950s there was "no evidence for a genetic factor"; now, as they

show, it is again considered well established. Similar U-shaped patterns can be shown for the conclusions of medical science about the role of alcohol in cirrhosis, and about the unhealthy effects of tobacco smoking, while Erickson mentions a reverse U — with about the same timing — of opinion about the medical effects of heroin use. The U-shaped curves concerning alcohol can be seen as direct reflections of the cycles of societal response to drinking: the first peak of the curve represents the zenith of the classical temperance movement in the early years of this century, while the second peak corresponds to the new wave of societal concern about alcohol, which has gradually built over the last 40 years. With respect to the U-shaped curve of scientific *output,* such an interpretation is not particularly threatening to the ideal of science as an objective pursuit of knowledge. A societal investment in research on addiction has mostly been driven not by a generalized commitment to expanding knowledge but by a desire for practical knowledge in dealing with a social problem. In light of this, it is not surprising that the scientific output ebbs and flows, in part, in tune with ebbs and flows in societal attention to a problem.

THAT the scientific *conclusions* seem to change back and forth according to changes in the general cultural position of drinking or drug use is much more threatening to the idea of science as an objective rather than culture-bound exercise. A closer examination of the history, however, suggests that we are not dealing with a simple oscillation, but rather with a helix — in fact, with a kind of Hegelian dialectic. Nothing will better make the name of a scientific 'young Turk' than proving faulty the methods or results of a received piece of work. Temperance science, often carried on with more commitment than training, certainly had its faults, and expert critiques of it in the first decades of this century did much to undo its legacy. The temperance-era studies of the inheritance of alcoholism, for instance, were demolished by Francis Galton and his students as part of a general upgrading of statistical work. Raymond Pearl, among others, did similar damage to temperance-era insurance company studies of drinking and mortality. And John Koren undercut the temperance-era assertion that three-quarters of all crime was due to alcohol. These demolition jobs occurred in, and were part of, a period of societal 'deflation' of alcohol problems. They were followed by a long fallow period during which few new studies were done, and the issues of the inheritance and the consequences of drinking were regarded

as largely settled. But as North Americans began to take an interest in alcohol problems again after the Second World War, and as new studies were undertaken, the received wisdom of the 1930s and 1940s began to come under attack from a new generation of young Turks.

MEANWHILE, general advances in scientific methods in the half-century or more since the heyday of temperance science meant that researchers setting out in the 1950s felt that they had an unusually open and fruitful field in which to work. From the lectures in this book, we can get some sense of the progress they have been able to make over the last 40 years. The general directions of some of the conclusions may resemble those of the temperance era, but there is no comparison in the past for the level of detail and the quality of evidence that we can take for granted in the best present-day scientific work.

THIS IS NOT to say that there will not be further turns of the helix as societal attention changes with respect to alcohol and drug issues. New perceptions and new techniques will undoubtedly present opportunities to chip at weak points in our current understandings of alcohol and other drugs. The work of these 40 years, however, will leave a permanent mark. This will be most obvious, of course, in future scientific work. But it will also be apparent on future societal policies on alcohol and other drugs. Scientific knowledge does not often determine the shape of public policies. But at a minimum, scientific findings do set limits to the public discourse. And in the long run, they may also have a more direct influence on policy. The world will be better served if the next 40 years of work at the Addiction Research Foundation and elsewhere are as influential both in the world of science and in that of policy as has been the work of the past 40 years.

Robin Room
Vice-president, Research and Development
Addiction Research Foundation
Toronto, Ontario

ACKNOWLEDGEMENTS

IN 1989, the Addiction Research Foundation celebrated its 40th anniversary. A major event of that year was the scientific lecture series. A group of international scholars addressed major developments in their fields over the past 40 years. These lectures were presented between February and November, 1989, but authors were given the opportunity to revise and update their material before this publication of their papers.

THE TASK OF bringing the 40th anniversary scientific lecture series to fruition as a book has drawn on the talents of several individuals within ARF. Special thanks are due to Jocelyne Soulodre and Denise Rochwerg for their energy and commitment in guiding the publication of this collection, to Nancy Leung for the attractive artwork and design, to Howard Shrier for able editing services and to Barbara Shimizu for expert secretarial support. Elda Hauschildt and Anne MacLennan kindly made available photographs of the authors from *The Journal* files. The editors are also grateful to Robin Room for graciously agreeing to provide the Foreword so soon after his arrival at the Foundation. These contributions have helped to make this commemorative volume of distinguished scholarship, celebrating the Addiction Research Foundation's 40 years of excellence in scientific research, available to a wider audience.

P.G.E.
H.K.
Toronto, Ontario
May, 1992

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

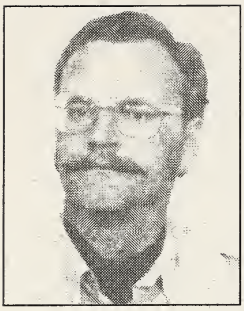
Epidemiological and Biostatistical Aspects of Alcohol Use, Alcoholism, and their Complications

OLE-JØRGEN SKOG

NATIONAL INSTITUTE FOR

ALCOHOL AND DRUG RESEARCH

OSLO, NORWAY



OLE-JØRGEN SKOG

A graduate of the University of Oslo, Ole-Jørgen Skog began his career in the addictions field as a research assistant at Oslo's National Institute for Alcohol and Drug Research in 1970. Three years later, he became a scientist at the Institute, and in 1988 was named director. He continues to hold that position while also teaching epidemiology and biostatistics at the Kongsborg College of Engineering. His recent work centres on the epidemiology of drug use and HIV/AIDS among drug users; on suicide and alcohol; drinking careers; and methodological issues, including the philosophy of science. Recent publications include: "The validity of self-reported drug use," *British Journal of Addiction* (in press); "The development of alcoholics' and heavy drinkers' consumption: A longitudinal study," *Journal of Studies on Alcohol* (in press); and "Alcohol and suicide — Durkheim revisited," *Acta Sociologica*, 1991.

1. INTRODUCTION

One of the basic tasks in alcohol epidemiology is to describe and explain variations in population rates of alcohol abuse and alcohol-related problems, within each culture over time, as well as cross-culturally and between population groups within each culture. In this article, my aim is to discuss some of the determinants of these aggregate problem rates, but I shall also discuss a few individual level issues whenever necessary in order to understand aggregate level phenomena.

The concept of risk is a useful starting point for this endeavor, since it can serve the purpose of tying together different key issues, and allows an explicit statement of relevant assumptions. Hence, I start with a discussion of dose-response relations or risk functions for different types of damage, and the aggregate level patterns to be expected in different cases (section 2). It is demonstrated that for some types of damage, the distribution of alcohol consumption in the population strongly affects aggregate level problem rates.

This issue is discussed in section 3, where the Ledermann theory and empirical results concerning the distribution of alcohol consumption are briefly reviewed. Section 4 is devoted to theoretical aspects of the distribution and the task of explaining the regularities observed.

Chronic diseases such as cirrhosis of the liver and cancer of the esophagus require long-term abuse. In section 5, I discuss the time factor and its implications for epidemiology of chronic diseases, particularly the time-lag between drinking and damage. In section 6, a related issue is discussed, namely the effect of stability of individual drinking careers on aggregate problem rates. Recent trends in liver cirrhosis mortality are briefly reviewed in section 7, in an attempt to bring together the impact of the different factors discussed in previous sections.

Section 8 is devoted to acute problems, and discusses some special issues in relation to this class of alcohol problems. Prevention implications are discussed in section 9.

2. RISK FUNCTIONS

A central theme in alcohol research has been to identify how the risk for different types of detrimental effects varies with level of intake. At what level can most people safely drink without running any significant risk? And how does the risk increase when intake increases?

Besides being of interest in its own right, the study of risk functions represents an important link in the study of problem rates at the aggregate level. The shape of the risk function to a large extent determines how aggregate problem rates vary in response to changes in population drinking habits. I shall illustrate this with some examples, focusing on three major cases (**Figure 1**).

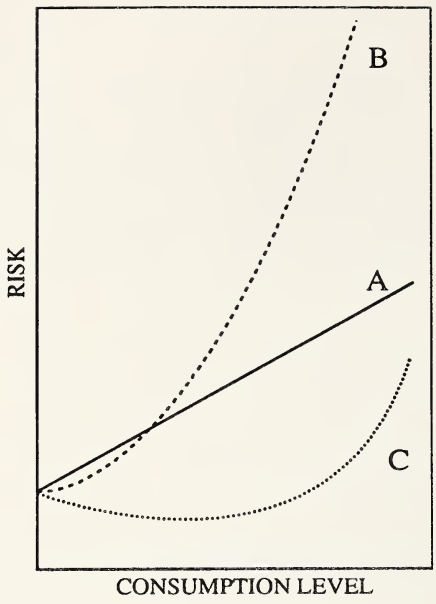
Figure 1.

Hypothetical risk functions

Linear (A)

Convex (B)

U-shaped (C)



A. Linear risk function

The simplest case occurs when the risk for a certain disease increases linearly with intake. With zero intake the risk is at some baseline level, and the risk increases with a fixed quantity for each unit by which intake is increased.

In this case it is easily shown (cf. **APPENDIX 1**) that the relationship between consumption and risk generalizes to the aggregate level without further qualification. Keeping other factors fixed, the prevalence rate in the population becomes a linear function of the average consumption level in the population. Exactly how alcohol consumption is distributed in the population is immaterial in this linear case. Populations with the same mean consumption level, but very different distributions, will have the same morbidity and mortality rates provided that everything else remains the same (cf. **Table 1**, disease A).

Therefore, in the case of diseases with linear risk functions, a change in the population drinking level will have predictable effects on mortality and morbidity. It is immaterial whether the increase has produced more heavy drinkers or not.

B. Convex (positively curved) risk function

However, for most chronic diseases the risk function seems to be non-linear. Both cirrhosis of the liver (Pequignot *et al.*, 1978), cancer of the esophagus (Tuyns *et al.*, 1977) and pancreatitis (Sarles, 1985) have been shown to have convex risk functions, increasing nearly exponentially with intake. In such cases the aggregate level prevalence rate for the disease will strongly depend on how alcohol

Table 1. Mortality for two hypothetical diseases with linear (A) and exponential (B) risk functions, respectively, in three model populations with identical per capita consumption, but different distributions.

Consumption level	Risk per 100,000 for disease		POPULATION I			POPULATION II			POPULATION III		
			No. of persons in 1,000	Expected number of deaths		No. of persons in 1,000	Expected number of deaths		No. of persons in 1,000	Expected number of deaths	
	A	B		A	B		A	B		A	B
0	1	1	20	.20	.20	20	.20	.20	20	.20	.20
1	2	4	40	.80	1.60	55	1.10	2.20	58	1.16	2.32
2	3	16	40	1.20	6.40	15	.45	2.40	11	.33	1.76
3	4	64				6	.24	3.84	6	.24	3.84
4	5	256				3	.15	7.68	3	.15	7.68
5	6	1024				1	.06	10.24	2	.12	20.48
Mortality per 100,000				2.20	8.20		2.20	26.56		2.20	36.28

consumption is distributed in the population. Three populations with identical mean consumption levels, but different distributions, may therefore have highly different mortality rates (cf. **Table 1**, disease B).

In this case one may consequently expect a close relationship between the aggregate consumption level of alcoholic beverages and morbidity and mortality rates, only if there are certain regularities in the distribution of alcohol consumption in the population. The more curved the risk function, the stronger the regularity required to secure a strong relationship at the aggregate level. Thus, in this case an increase in population drinking may or may not produce an increase in morbidity and mortality, depending on whether or not the increase has produced more heavy drinkers.

When the risk function is very strongly curved, as appears to be the case for the diseases mentioned, even quite minute changes in the distribution of alcohol consumption may have dramatic effects on mortality rates. This is illustrated in **Table 1**. Populations II and III in this hypothetical example have identical mean consumption levels, and both distributions are very skewed towards a high consumption level — population III somewhat more than population II. This minor difference in the tails of the distributions — too small to be detected in an ordinary survey — has a very large impact on mortality rates. The reason for this state of affairs is that an exponential risk function implies that even quite moderate changes in the consumption level of a heavy drinker or an alcoholic — say, a decrease from 120 to 100 g/day —

reduces his risk for cirrhosis of the liver by more than 50 percent, according to the data published by Pequignot and Tuyns (1978). In fact, the relative risk would drop from 92 to 43 as a result of this decrease in consumption (cf. Skog 1985b).

C. U-shaped risk function

Several studies have suggested a U-shaped risk function for ischemic heart disease in relation to alcohol (Williams, Robinson and Balley, 1979; Yano, Rhodes and Kagan, 1977). Moderate amounts of alcohol have thus been hypothesized to offer some protection against this disease due to favorable effects on the HDL cholesterol (Johansson and Medhus, 1974). However, most of these studies suffer from serious methodological shortcomings. The measurement of alcohol consumption is problematic in many of these studies (Knupfer, 1987), and confounding factors (e.g. nutrition) are often not adequately controlled for (Popham, Schmidt and Israel, 1983). The status of the 'protection' hypothesis therefore remains somewhat obscure. However, in spite of all this, it is of interest to take a brief look at the aggregation problem for U-shaped risk functions.

When the risk function is U-shaped, there is an optimal consumption level where the risk is minimal. In the case of ischemic heart disease, 10 to 20 g/day has been suggested. At the aggregate level, a U-shaped risk function would produce a U-shaped relationship between mean consumption and mortality rates, provided that there is sufficient regularity in the distribution of alcohol consumption. Hence, one can infer the existence of an optimal per capita consumption level as well. At this consumption level mortality is at a minimum. However, it is important to observe the following two facts.

Firstly, the potential reduction in the population morbidity and mortality rate is much less than the risk reduction for an individual drinking at the optimal level — provided that the population consumption level is optimal. This is simply due to the fact that a population is heterogeneous, and there will always be people who drink too much or too little, compared to the optimal level. The reduction in mortality rate in a population drinking at the optimal level could easily be less than the optimal individual reduction in relative risk by a factor of three to five (cf. **APPENDIX 1**). Hence, if the individual risk is reduced by maximum 10 per cent, the decrease in population mortality may be only two or three per cent.

Secondly, the optimal level of intake would typically be much smaller for the population than for the individual. This is because people tend to change their drinking by roughly the same percentage when changes occur (cf. section 2), i.e. multiplicatively rather than additively. Due to this effect the population optimum consumption level may easily be smaller than the individual optimum level by a factor of three to five (cf. **APPENDIX 1**), which would mean that the optimal mean

consumption level for a population would be between one and three litres of ethanol per year. This is much less than is consumed in most Western countries, and it may therefore very well be the case that all these countries are all beyond the limit where they would profit from the alleged protective effect.

3. THE DISTRIBUTION OF ALCOHOL CONSUMPTION – EMPIRICAL RESULTS

As the considerations above exemplify, the aggregate morbidity and mortality rates depend on the shape of risk functions, as well as the distribution of alcohol consumption in the population. Provided that there are regularities in the distribution, across cultures and within each culture over time, one would expect covariation between morbidity and mortality rates and the overall level of consumption in the population.

The question of regularities in the distribution pattern has received much attention in the literature over the past 25 years, since ARF researchers first brought the issue to the attention of the international scientific community (deLint and Schmidt, 1968; Schmidt, 1975). The stepping stone for this line of research was Ledermann's so-called single distribution theory (Ledermann, 1956; 1964), which hypothesized very strong regularities in the distribution pattern. Ledermann suggested that alcohol consumption followed a one-parametric variant of the lognormal distribution, to the effect that the complete distribution is specified once the mean consumption level per drinker is known.

The Ledermann theory in its original form has been rightfully questioned, both on theoretical and empirical grounds (cf. Duffy, 1986; Skog, 1982). Empirical distributions often deviate significantly from the lognormal model, and the relationship between the parameters is less fixed than Ledermann predicted. However, Ledermann was right in his basic presumption, namely that substantial regularities do exist. The basic findings can be summed up as follows (cf. Skog, 1985):

1. Empirical distributions are typically strongly skewed with a long tail towards high consumption levels. In most populations every seventh drinker (i.e. about 14 per cent) consumes more than twice the average level in the population.

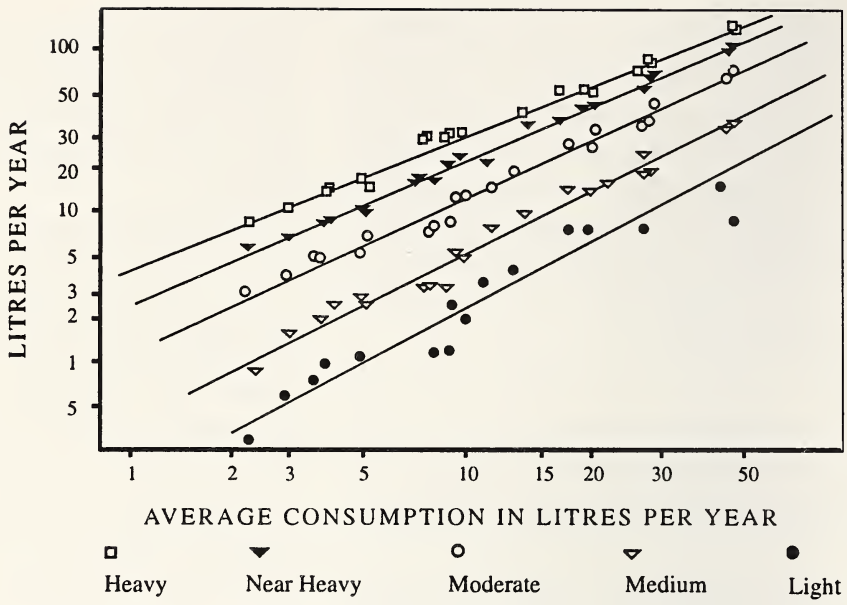
2. The standard deviation of the distribution typically increases strongly when the average increases. The relative dispersion, measured either by the coefficient of variation or the standard deviation of logarithmic consumption, tends to decrease slightly when average consumption increases.

3. No sharp line of division between alcoholics and normal drinkers has been found.

4. There are fairly close log-linear relationships between the average consumption and the different percentiles (cf. Figure 2). This means that changes (or

Figure 2.

Consumption level among different types of drinkers as a function of the mean consumption level in their population (Source: Skog, 1985).



differences between populations) in mean consumption are typically the result of parallel changes (or differences) in consumption among all types of drinkers. The population thus tends to move collectively up and down the consumption scale when the mean changes.

5. As a result of this, there is a fairly close relationship between the average consumption level and the prevalence of drinkers at high levels of intake, say, 40 or 80 g/day or more (**Figure 3**). Typically, the increase in the prevalence of heavy drinkers is disproportionately large, compared to the increase in average consumption, and as a rule of thumb a doubling of the average consumption is accompanied by a fourfold increase in prevalence of drinkers whose daily intake exceeds 80 g/day.

These results are based on comparisons of survey data from different countries, but the data base now also includes an increasing number of repeated surveys from the same culture (cf. Skog, 1985e, Knibbe *et al.*, 1985). The results obtained in cross-cultural comparisons are typically confirmed when repeated measurements are available.

Although a substantial regularity in the distribution pattern and the pattern of change has been demonstrated, one should not forget that these regularities are statistical in nature and that exceptions are likely to exist.

One such exception has been demonstrated by Norström (1987a) in data from Sweden. Comparing both survey data and purchasing data collected prior to and after a major reform in Swedish alcohol control policies, a substantial change in the distribution pattern was uncovered. Prior to the

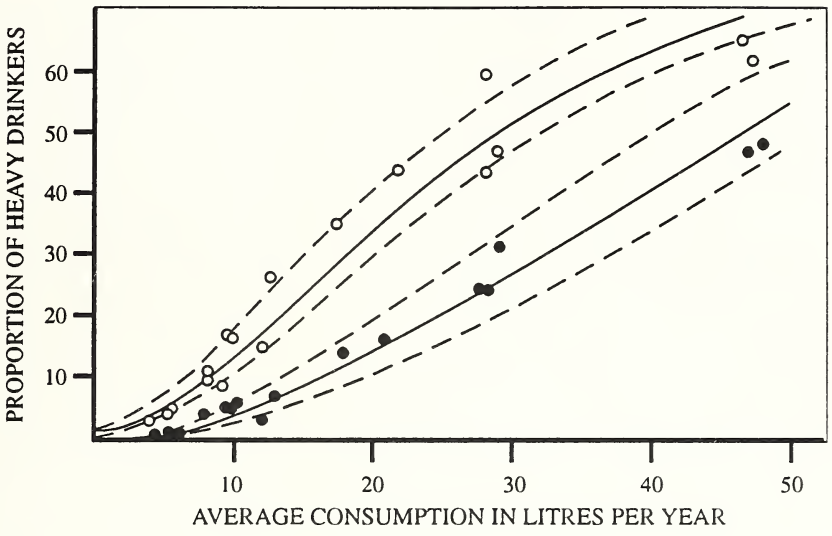


Figure 3. Relationship between mean consumption per drinker and the prevalence of drinkers whose annual intake exceeds 5 (open circles) and 10 cl (closed circles), respectively. The solid curves are least square regression curves and the dotted curves are one standard error confidence limits (Source: Skog, 1985).

reform, the variance of the distribution was much smaller than could be expected on the basis of data from other countries, while the variance was more 'normal' after the reform. The per capita consumption level did not change to any significant extent during this period. However, problem indicators — both cirrhosis mortality rates (Norström, 1987b) and crimes of violence (Lenke, 1983) — increased substantially along with the increase in the variance.

The reform in question was quite substantial — namely the abolition of a rationing system that had been in operation for about 40 years. *A posteriori*, it may perhaps not appear very surprising that rationing strongly suppresses the drinking of heavy and near-heavy drinkers, and thus serves to shrink the tail of the distribution. However, *a priori*, many would probably have argued that rationing is not likely to be effective, since a black market would develop. No doubt, a large black market existed in Sweden during the rationing era, but since prices were typically much higher on the black market, the result really was a two-step price system. Moderate amounts could be bought at a low price, while for additional amounts the price was much higher.

This Swedish example should serve as a reminder that the regularities observed in the distribution of alcohol consumption are neither laws of nature nor iron clad rules. The regularities are the result of an interplay between many different factors. Large changes in some of these factors may produce significant departures from the main pattern. Of particular importance are socio-cultural and political factors. When strict regulations on individuals'

drinking prevail, as in the Swedish case, exceptional patterns ought to be expected. The results of the research on the distribution of alcohol consumption can be said to have paved the way for a general presumption that changes in total consumption are typically the result of collective changes and accompanied by similar changes in the prevalence of heavy drinking. However, in individual cases, and under special circumstances, things may look different. In order to better understand when deviations from the 'rule' are likely to occur, a better understanding of the processes shaping and reshaping individual drinking habits are needed. In short, one needs an explanation for the regularities observed.

4. SOCIAL INTEGRATION AND THE VARIANCE OF THE DISTRIBUTION

In its original form, the Ledermann theory was a purely statistical generalization, and the theory did not offer a substantial explanation for the regularities it predicted.

In the so-called 'social interaction theory' of the distribution of alcohol consumption (Skog, 1979; 1980), a theoretical re-specification of the Ledermann theory was attempted. The social interaction theory takes as its point of departure some basic social and bio-psychological mechanisms. The main postulates of this theory are: (1) that an individual's drinking is influenced by numerous social, psychological, and biological factors that tend to combine multiplicatively, rather than additively; and (2) that individual drinkers of all types tend to influence each other's drinking habits mutually, both directly and indirectly.

These propositions are well supported both in observational data and experimental studies, both of normal drinkers and alcoholics (cf. Skog, 1980a, for an overview of the literature). The first postulate, which is often called the "law of proportionate effects," is related to Weber-Fechner's law in psycho-physics, and implies that the distribution of alcohol consumption is skewed and roughly resembling the lognormal distribution. The second postulate will, due to what I have called the social network "principle of long-range indirect ties" (Skog, 1986b), link together individual level and aggregate level changes.

In fact, it can be demonstrated that mutual influences between individual drinkers tend to produce an overall synchronization of individual changes within a culture (Skog, 1979; 1985a; 1986b). Hence, changes in the consumption of alcoholic beverages tend to have a significant collective component, and individual drinkers should tend to move up and down the consumption scale more or less in concert, when per capita consumption changes. By the same argument, differences between cultures with respect to alcohol consumption tend to be

collective differences, and going from one culture to another, one should observe a shift of the whole distribution of alcohol consumption.

The social interaction theory thus stresses the social fabric of drinking habits in general, and of heavy drinking in particular. Informal social control and learning by social modelling are crucial concepts. Hence, there are important links between the social learning theory approach to alcoholism (Marlatt and Nathan, 1978; Miller, 1976) and the social interaction theory of the distribution of alcohol consumption.

Numerous studies have demonstrated that the physical, legal and economic availability of alcoholic beverages very often have an impact on consumption (cf. Bruun *et al.*, 1975, for an overview). The collective patterns observed in distribution data point to the importance of availability in a cognitive and cultural sense as well. The more widespread the drinking of alcoholic beverages is as a culturally meaningful act, and the larger the number of contexts in daily life where this occurs, the more natural this type of behavior becomes for the individual, and the larger the risk that an individual shall reach consumption levels that threaten his health.

The social interaction theory predicts regularities in the distribution pattern and in patterns of change. However, the theory also suggests under what circumstances deviations are to be expected, by pointing to factors affecting the variance of the distribution (Skog, 1979; 1980a). Two factors of particular importance are the strength of informal social control, and the connectedness of social networks.

Strength of informal social control means the degree to which individuals influence each other's drinking. In groups with strong informal control, i.e. in strongly integrated groups, intra-group variability ought to be small, as compared to weakly integrated groups. Excessive drinking would therefore be expected to occur more frequently in weakly integrated groups.

Society at large can be viewed as an enormous social network of interconnected individuals. Connectedness is a structural property, determining how easily diffusion processes penetrate the network. It can be shown analytically that the larger the connectedness of a network, the smaller the variance becomes, and the smaller the number of excessive drinkers, provided everything else remains constant (Skog, 1979).

Both of these factors are aspects of the more general sociological concept of social integration. We are thus led to a hypothesis that the variance of the distribution of alcohol consumption is inversely related to the degree of social integration of society — or solidarity in Durkheimian terminology. (This statement should not be confused with the "integration of drinking" hypothesis of the socio-cultural models, cf. section 8.)

This line of reasoning suggests that some connections may be made between

the theory of the distribution of alcohol consumption and Durkheim's theory of suicide (Durkheim, 1951; Skog, 1991). Durkheim explains variations in suicide rates, both between different substrata and between countries, in terms of variations in integration in society. In particular, what Durkheim calls "egoistic suicide" is hypothesized to increase with decreasing integration. This suggests that one possible causal path in the link between alcohol abuse and suicide may be indirect, mediated by social integration. To put it differently, a low level of social integration may produce both a high level of abuse in society, and a high suicide rate.

To my knowledge, no large-scale empirical documentation of the alleged relationship between level of integration and rates of abuse is available. It is worth noting however, that a study by Schmidt and Bronetto (1962) produced evidence that the degree of urbanization correlates positively with cirrhosis mortality rates, after controlling for per capita consumption. Urban society thus seems to be less effective in controlling deviations from the mean, and it does not appear unlikely that this is in fact the result of differences in level of social integration. This conjecture is also corroborated by Gove and Hughes (1980), who demonstrated that both suicide rates and alcoholism rates are higher in regions that have larger numbers of single-person households.

At present, this is not much more than armchair speculation. However, the basic hypothesis — namely that the level of social integration may affect the variance of the distribution and thus the prevalence of excessive drinking in society — can be tested empirically. I would argue that this is an important line of research, particularly because it would create some links between alcohol epidemiology and general sociological theory.

5. THE TIME FACTOR

Besides being strongly dependent on consumption level, it is well known that the risk for many types of detrimental effects, such as cirrhosis of the liver, strongly depends on the duration of alcohol abuse (cf. Lelbach, 1976). In order to better understand the epidemiology of alcohol-related chronic problems, the time factor clearly needs to taken into consideration.

The problem of duration should not be addressed, however, by asking how long it takes to develop cirrhosis of the liver. For one thing, this question is not meaningful unless one specifies which consumption level one is talking about. At moderately high consumption levels it obviously may take a very long time, while at very high levels much less time would be needed. Secondly, even when the consumption level is specified, the time needed may typically vary tremendously from one individual to the next, and the average duration may not be a very meaningful measure. Therefore, one

should instead ask how the risk for cirrhosis increases with duration of abuse, or more precisely how the risk depends on the person's drinking career.

A person's alcohol abuse may typically vary a great deal during his drinking career. Periods of abstinence, say due to treatment, and relapses to drinking are typical events in a heavy drinker's career. This makes it difficult to describe the temporal aspect by a single summary measure of the duration of abuse. Clearly, the risk of cirrhosis of the liver depends on the present drinking level as well as the drinking level during successive periods in the past. We are thus faced with the multivariate problem of describing how the risk is jointly determined by the present level of intake, as well as past levels of intake.

The best way of approaching this issue is probably to start from the presumption that the risk at any particular point in a person's life is causally determined by his cumulated drinking experiences up to that point, and that a reasonable measure of a person's cumulated drinking experiences is the weighted sum of present and past drinking throughout the whole drinking career. In this formulation, the weights measure the etiological importance of past drinking, and the distribution of the weights gives a description of the etiological process. According to available data, the weights measuring the importance of past drinking decrease as we move backwards in time (Skog, 1985c). Therefore, the drinking career measure of lifetime exposure is not simply equal to the total intake during the person's lifetime — a measure used by Lelbach (1976). While the latter places equal weight on each year of the past, the former measure attributes more weight to present than to past drinking.

At the aggregate level, this implies that present mortality is causally determined by present, as well as previous, drinking. Hence, the relationship between consumption level and mortality rate has a so-called distributed time-lag structure. It implies that changes in mortality often tend to follow a fairly complex pattern, and that the effects of shifts in consumption trends may not always be simple and straightforward.

The time-lag is frequently misinterpreted as a delayed response, i.e. a tendency for mortality rates to 'wait' a number of years before responding to a change in consumption. However, the pattern to be expected on theoretical grounds, and actually observed in the data (Skog, 1985c; Norström, 1987a), is an immediate but gradual response. The effect of a change in consumption typically becomes partially visible more or less immediately, but the full effect builds up gradually over a period of many years (cf. Skog, 1980b).

When studying trends in chronic alcohol-related problems, such as liver cirrhosis, it is necessary to take into consideration this time-lag between consumption and the disease. If one simply correlates present consumption

level in a population with present mortality rate, one effectively ignores the fact that it normally takes considerable time to develop cirrhosis of the liver. In some cases very wrong results may therefore be obtained.

An example of this, originally reported by Bronetto (1961) and later quoted by Popham (1970) in his well-known overview, is the so-called 'British exception.' During a large part of the period 1931-58, alcohol consumption increased in the U.K., while cirrhosis mortality decreased, producing a negative concurrent correlation of - 0.61. It turned out, however, that this was a time-lag effect, rather than a signal of a breakdown in the alleged relationship between per capita consumption and prevalence of heavy use (Skog, 1980b). Alcohol consumption had been decreasing in the U.K. during the first third of this century. This trend was reversed in 1931, and consumption started to rise again, although slowly. Mortality, however, continued to decline for another ten years or so before starting to rise again. Since about 1950, the two trends have again been moving in concert.

This tendency for mortality continuing to decrease after consumption has stopped decreasing is typical when there is a strongly distributed lag. The length of this transition period is not fixed, but will depend on several things. First of all, there is the nature and strength of the time-lag, i.e. how the parameters measuring the relative importance of successive years in the past decline backwards in time. Secondly, we must consider the nature of the consumption trend prior to and after the turning point. The longer the downward trend has lasted, the more time is needed before the system will readjust itself. In fact, there is no definite answer to the question of how much time it takes before a trend-reversal in consumption manifests itself in the mortality data.

Technically, the time-lag can be taken into consideration by seeing mortality rates for any specific year in relation to a cumulative, aggregate consumption measure, analogous to the individual drinking career measure. This cumulative aggregate measure is thus a weighted sum of per capita consumption in the same year and preceding years. In the case of cirrhosis we now have a reasonably good idea about the distribution of these weights, since they have been estimated in several studies, using different methods and different data sets, with similar results (cf. Skog, 1980b; 1985c; Norström, 1987a). Hence, it is now possible to make predictions about mortality trends on the basis of consumption trends that take the time-lag into consideration. Thus it seems possible to approach the issue of deciding if current trends in mortality conform to expectations on the basis of consumption trends and the distribution theory, or if discrepancies are present. We shall return to this issue below.

6. DRINKING CAREERS AND CHRONIC DISEASES

Students of population trends and cross-cultural differences in chronic health problems due to alcohol have shown surprisingly little concern with the nature of drinking careers, and the impact of changes or cross-cultural differences in careers on mortality rates.

Two populations may have the same average consumption level and identical distributions of annual intake, but still be very different in terms of stability of individual drinking. This obviously would have an impact on mortality and morbidity for diseases requiring long-term abuse. In a culture where abusers are very stable in their abuse, detrimental effects ought to be greater than in a culture with less stable patterns. In the latter case, turnover rates would be high, with many people going in and out of abuse each year, while in the former case, the abusers would be the same individuals each year. Therefore, although the distributions of annual intake are identical, the distributions of long-term intake would be very different, and the prevalence of long-term heavy drinkers should be higher in a stable culture.

Some rough estimates of the quantitative effect of variations in career stability can be obtained (**APPENDIX 2**). It turns out that the effect may be quite substantial for diseases with convex risk functions. Hence, if cultures are different in terms of stability of drinking careers, or if changes in stability should occur, substantial differences in mortality rates that are not explained by differences in per capita consumption may occur.

Herd (1985) has suggested that the differences between white and black liver cirrhosis mortality in the United States may to some extent be related to such differences. However, her hypothesis is not based on data showing actual differences in stability of abuse, but is more indirect and based on differences in the age distributions of heavy drinking among black and white people. The hypothesis is of considerable interest, however, and the matter deserves to be looked into more closely.

Some authors have suggested that the expansion in the treatment sector in many countries may have had effects on morbidity and mortality (Romelsjö, 1987; Mann *et al.*, 1988). Although the experimental evidence that treatment has a significant and lasting effect on abuse are not entirely convincing, it may still have effects on morbidity and mortality via a mechanism of the type outlined above. Clearly, treatment may contribute to intermittence of drinking, and thereby reduce correlations and risks.

In any case, it should be obvious that epidemiologists need to consider careers more closely than they have done in the past. In particular, a better understanding of the quantitative aspects of drinking careers is needed.

Considering the resources used on alcohol research during the last decades, we still know surprising little about natural

drinking careers among both users and abusers. This is probably so for at least three different reasons. Firstly, most studies seem to have a very small number of follow-ups of each individual — typically two or three. Hence, it becomes difficult to disentangle measurement error and real change. Secondly, measures of drinking and highly different consequences of drinking are often lumped together into dichotomous categories — often of a normative nature. One often fails to take advantage of the fact that we can measure several conceptually distinct and highly relevant dimensions on ratio scales. Thirdly, alcohol research suffers from a methodological and conceptual poverty when it comes to the problem of describing and analysing processes of change and temporal structure. The theory of stochastic processes as a method for describing and analysing dynamic processes is apparently not part of the statistical toolbox of most medical and social students of alcohol.

Two main schools of thought can be identified on the career issue. On the one hand, adherents to the medical model conceptualize alcoholism as a progressive disease (Jellinek, 1960), and it is normally assumed that changes in alcoholics' drinking typically have a definite direction. Furthermore, the disease concept supports a notion of inertia and resistance to change, which implies a fairly strong stability in the drinking habits of alcoholics, once the disease is established. In short, alcoholism is conceived as chronic. This conception is mostly based on observations of clinical populations, and the fact that many diagnosed alcoholics continue to drink heavily for many years.

On the other hand, it has been suggested on the basis of general population panels that 'problem drinking' is typically a highly unstable state with high turnover rates (Cahalan and Room, 1974; Clark, 1976; Clark and Cahalan, 1976). These and similar results have inspired a radically different picture of alcoholism, as a highly transient rather than stable state.

As Room (1977) puts it, these different approaches seem to face us with "two worlds of alcohol problems." Vaillant (1983) tries to resolve the apparent conflict between these conceptions by pointing to the fact that the two approaches focus on highly different groups of problem drinkers. The problem drinkers encountered in population surveys are typically much younger and less marginalized than those encountered in clinics, etc. A recent study by Fillmore and Midanik (1984) illuminates this facet. The authors demonstrate that chronicity is a function of age, and that the turnover rates are lower in the 40-49 age group than in the 20-29 age group. The authors suggest that the very high turnover rates reported in some studies may be due to the fact that "problem drinking" has often been defined in terms of harmful events, rather than in terms of more persistent or chronic problems.

Still, the resolution of the conflict is only partial. Fillmore and Midanik's data

show substantial instability even in the older age group, and Polich *et al.* (1981) also report a substantial instability among treated alcoholics.

The stability problem is at the very heart of the so-called 'return to normal drinking' controversy. Critics of the disease model suggest that spontaneous remission may be a fairly common phenomenon (e.g. Heather and Robinson, 1983). On the other hand, advocates of the disease model and related perspectives argue that return to normal drinking is not well documented, except perhaps in a very small fraction of cases. (cf. Vaillant, 1983). Many studies, however, suffer from methodological shortcomings, and as Edwards (1985) has recently shown, the evidence presented in Davies' classic study (Davies, 1985) is not very convincing. Edwards argues that only one or possibly two of the seven cases really had returned to 'normal' drinking.

However, less fragile evidence does exist for spontaneous remission. Drew (1968), studying incidence and prevalence rates as a function of age, demonstrates an increasing divergence between prevalence rates and cumulative incidence rates with increasing age. He argues that this is most likely a sign of a considerable spontaneous remission. Whether this is a remission to abstinence, 'normal' drinking, or 'less problematic' drinking is an open question.

Secondly, we know that the prevalence of heavy drinking co-varies quite closely with per capita consumption, and in certain historical periods the latter has decreased dramatically (cf. Bruun *et al.*, 1975). Also, problem rates (e.g. cirrhosis mortality) have declined rapidly and quite dramatically in certain periods (e.g. Skog, 1985b). These aggregate data must obviously imply that excessive drinking of the type leading to cirrhosis has decreased. Apparently then, significant changes towards less problematic drinking do occur, at least in connection with holocultural reductions in drinking.

In individual-level studies of 'return to normal drinking,' the remission rates strongly depend on how the concept 'normal' is defined, as well as the time frames. For instance, Roizen *et al.* (1978) have shown that the more stringent the definition of 'abuse,' the smaller the likelihood of spontaneous remission. And as Vaillant (1983: 220) argues, the shorter the time frames, the higher the chance that you will observe periods of 'normal' drinking.

In short, much of the variance in results concerning the stability of problem drinking may simply be a result of arbitrariness in definitions, and in the choices of cut-off points. Therefore, in order to progress it is necessary to develop strategies for studying natural drinking careers that are less dependent on such arbitrary methodological conventions. For instance, it may be a better strategy simply to record reductions and increases in consumption and objective consequences of drinking, without resorting to pre-fixed standards for 'normal' drinking. From the point of view of survival, a large reduction

in consumption may be quite important, even when the resulting drinking pattern does not qualify for the social label 'normal.'

7. RECENT TRENDS IN LIVER CIRRHOSIS MORTALITY

It is well known that trends in aggregate alcohol consumption and liver cirrhosis mortality co-varied quite closely in most Western countries until the early 1970s (Popham, 1970, Bruun *et al.*, 1975). In recent years, however, several countries have experienced downward trends in mortality, and these trends do not always parallel trends in consumption levels (Smart, 1988; Smart and Mann, 1987).

Over one recent 10-year period — from 1974 to 1984 — consumption remained practically unchanged (less than five per cent change) in nine of the 17 countries represented in **Table 2**. Consumption increased in four countries and decreased in four. The four countries with increasing consumption also experienced increasing cirrhosis mortality rates, and of the four countries with decreasing consumption, three experienced decreasing mortality. More intriguing is the fact that seven of the nine countries with stable consumption experienced declining mortality from liver cirrhosis.

From the mid-1960s to the mid-1970s, the pattern was more coherent. Of the 17 countries, only France experienced declining consumption, and at the same time a modest increase in liver cirrhosis mortality. Switzerland had a moderate increase in consumption, and a moderate decrease in mortality. However, the remaining 15 countries experienced increase in both consumption and mortality — and in some cases the increase was quite substantial.

Given the great impact that the per capita consumption theory (or the distribution of consumption theory) has had in recent years, it is only natural that the divergences observed in recent years produce speculations that the theory may not be correct after all. However, in the light of the foregoing it should be evident that the relationship between per capita consumption and mortality rates is far too complex to allow conclusions based on impressionistic evidence.

First of all, it is necessary to take the time-lag problem seriously, since divergent trends are indeed to be expected in periods with reversals in consumption trends. This can be done without too much difficulty, and it is surprising that no study of recent trends has actually done that — at least not to my knowledge.

Secondly, one needs to consider the fact that indicators of damage — such as the liver cirrhosis mortality rate — are influenced by several factors besides alcohol abuse. Diagnostic practices may change due to redefinitions in connection with, say, revisions of ICD (cf. Grant *et al.*, 1986), or for less obvious reasons

Table 2. Trends in alcohol consumption (litres of pure alcohol per capita) and liver cirrhosis mortality (per 100,000 population) in selected countries.

COUNTRY	ALCOHOL CONSUMPTION			LIVER CIRRHOSIS		
	1964	1974	1984	1964	1974	1984
Australia	6.88	9.31	9.37	4.9	8.3	7.6[1]
Austria	9.44	11.27	11.02	25.1	32.7	30.2[1]
Canada	5.20	8.22	7.98	6.4	11.6	8.8
Denmark	5.14	8.33	10.37	7.8	10.4	12.1
Finland	2.58	6.57	6.58	3.5	5.5	7.0
France	18.82	17.08	14.22	31.7	32.8	27.6[3]
Hungary	7.33	9.43	11.45	8.7	16.2	39.2[1]
Japan	3.95	5.34	5.87	10.1	13.4	14.0[2]
Netherlands	3.62	7.94	8.60	4.0	4.5	5.3[1]
New Zealand	5.40	7.83	8.06	2.8	5.4	3.8[1]
Norway	2.69	4.23	3.95	3.6	4.1	6.1
Poland	3.92	6.20	6.24	5.9	9.7	11.3
Portugal	13.21	13.72	12.86	29.3	31.3	30.1[2]
Sweden	4.43	6.18	5.17	5.7	10.5	8.2
Switzerland	10.42	11.18	11.12	15.4	14.8	12.3
USA	5.83	7.75	7.90	12.1	15.8	12.0[2]
West Germany	10.06	11.69	11.90	20.8	26.9	25.2[2]

[1] 1983 [2] 1982 [3] 1981

related to societal change (cf. Skog, 1985b). Furthermore, most diseases have multiple etiologies, alcohol being only one among several causes. If changes occur in these other causal factors — say, nutrition, environmental chemical agents or other diseases — the trend in the disease to be explained may not correspond to expectations based on trends in alcohol consumption (cf. Ipsen, 1950; Petersen *et al.*, 1986).

Thirdly, the stability of drinking careers — and particularly heavy drinking careers — may have changed. As was previously noted, such changes may have no effect on per capita consumption or the distribution of annual consumption, but could still have substantial effects on mortality rates. The expansions in the treatment industry may very well have had such effects, and Mann *et al.*'s analysis of Ontario data, which uncovered a significant correlation between changes in alcoholism treatment and cirrhosis discharge rates (Mann *et al.*, 1988), suggests that it may be worthwhile to look more closely at this possibility.

Last, but not least, the distribution of alcohol consumption may have changed. The theory of the distribution of alcohol consumption should not be interpreted dogmatically. The theory creates a general

presumption that changes in per capita consumption are accompanied by parallel changes in heavy drinking, but the relationship is not fixed, and under special circumstances the latter effect may fail to be obtained, or it may be smaller than expected. As was pointed out in section 2, the effect of minor changes in the tail of the distribution on mortality rates may be quite substantial.

Deaths due to liver cirrhosis form only a fraction of all deaths due to chronic abuse. Although trends in cirrhosis mortality are probably closely correlated with trends in other alcohol-related deaths from chronic abuse, the picture easily becomes somewhat distorted as long as cirrhosis is the only indicator used. Therefore, it is desirable to extend the monitoring of trends to include other diseases as well. However, the evaluation of such trends is probably even more difficult than in the case of cirrhosis. This is due to the fact that other etiological factors may vary quite substantially over time for diseases such as cancer of the esophagus, pancreatitis, etc., and that these variations in other etiological factors need to be controlled before the effects of alcohol can be evaluated. However, with the aid of modern statistical techniques, this is not an impossible task.

8. OTHER TYPES OF DAMAGE – OTHER RISK DIMENSIONS

So far, I have mainly concentrated on somatic problems due to long-term abuse. For this class of problems the quantity consumed is the main etiological variable.

In the case of problems due to acute intoxication it is no longer sufficient to consider only the amounts consumed. Other aspects need to be considered, namely drinking patterns (frequent drinking of small amounts vs. less frequent drinking of large amounts), the drinking contexts, and the cultural rules for drunken behavior. All these dimensions affect the risk for acute drinking problems, and would potentially also affect aggregate problem rates.

It is well known that patterns of drinking, as well as drunken comportment, vary enormously from one culture to the next (MacAndrew and Edgerton, 1969). Since, in the Western World, explosive drinking typically occurs in countries with a fairly modest per capita consumption level, one should not expect any close relationship between per capita consumption and rates of acute problems in cross-cultural comparisons. Thus, in view of the large differences that exist between Finnish and Danish drinking cultures, it is not surprising that alcohol poisoning is much more frequent in Finland than in Denmark (Poikolainen, 1977), in spite of the fact that the Danes drink more than the Finns.

In general, relationships between per capita consumption of alcohol and rates of acute problems seem to be fairly weak in cross-cultural comparisons (Mäkelä, 1978). However, as Schmidt and Popham (1978) have pointed out, this does not prevent close relationships within each culture over time. Patterns of drinking and cultural rules for drunken comportment seem to change much more slowly than the amounts consumed, and increasing (or decreasing) consumption may therefore create parallel changes in acute problems within each culture. It should be recognized, however, that the rate of increase in acute problem rates may be expected to vary considerably from one culture to the next, due to differences in changes with respect to social norms and patterns of drinking.

Several studies confirm the existence of intra-cultural correlations between per capita consumption and acute problems over time (cf. Mäkelä, 1978, for an overview). However, there are also numerous examples where such correlations do not occur. For instance, Mäkelä *et al.* (1981) have shown that during periods of increasing consumption, arrests for drunkenness have increased in some countries and decreased in others. The latter cases may not necessarily indicate that drunkenness has decreased parallel to the increase in consumption. It may simply be a result of changes in formal social control.

In fact, data from Norway (Skog, 1987) point in this direction. Arrests for drunkenness decreased while consumption increased. However, a close analysis of the data suggested that drunkenness actually had increased, while the likelihood of becoming arrested had decreased, most likely due to changes in police practices.

Another thing to be considered is the fact that alcohol is only one of several causal factors affecting rates of acute problems, such as accident rates and rates of crimes of violence. These other causal factors may change, possibly in a different direction than alcohol, and the net effect on problem rates may not correspond very well with trends in alcohol consumption. Hence, comparison of trends is not a safe method for evaluating the effects of changes in consumption level. This was exemplified in a study of accident mortality in Norway (Skog, 1986a), where a strong increase in alcohol-related violent deaths was 'compensated' by a decrease in non-alcohol-related violent deaths, particularly industrial accidents. Since the statistics do not single out alcohol-related deaths, this result was not available to the naked eye, but could be unveiled by special analysis.

It thus appears possible that many of the negative cases reported in the literature, where problem rates have not followed consumption trends, may be rather poor evidence for a lack of correlation between consumption and alcohol-related problem rates. However, by the same argument, many positive cases may also be flawed, since coinciding trends are not necessarily good

evidence for a causal relationship.

The multiplicity of factors influencing rates of acute problems clearly indicate that an extended theoretical framework, where both quantitative and qualitative aspects of the drinking culture are integrated, is needed in relation to acute and more indirect problems. Such a model has been suggested by Whitehead and Harvey (1974), and further developed by Frankel and Whitehead (1981), who have tried to integrate the distribution of consumption theory with the socio-cultural theories suggested by Ullman (1958). Frankel and Whitehead (1981) explain the extent of alcohol-related damage in a society in terms of per capita consumption level, integration of drinking practices in everyday life, and the social norms regulating drinking: prescriptions for moderate drinking and proscriptions against excessive drinking. Analysing cross-cultural data from 68 societies, the authors find that all these factors have explanatory power and interact in a fairly complicated way to determine the end result.

However, it should be remembered that Frankel and Whitehead's data are not ideal. In particular, their measure of 'problem rate' is an overall summary measure where all kinds of damage are indiscriminately included. It would be of considerable interest to see a similar analysis done separately for each of the main types of alcohol-related problems. It would not be very surprising if the relative importance of quantitative and qualitative factors were very different for chronic and acute problems. One would also expect different results, particularly in relation to acute problems, from cross-cultural and temporal data.

9. IMPLICATIONS FOR PREVENTION

Numerous studies have demonstrated that the physical, legal and economical availability of alcoholic beverages often have an impact on consumption (cf. Bruun *et al.*, 1975 for an overview). The collective patterns observed in the distribution of alcohol consumption point to the importance of availability in yet another sense, namely in a cognitive and cultural sense. The more widespread the drinking of alcoholic beverages is as a culturally meaningful act, and the larger the number of contexts in daily life where this occurs, the more natural this type of behavior becomes for the individual, and the larger the risk that an individual shall reach consumption levels that threaten his health.

Since the prevalence of heavy drinking and chronic abuse is so intimately related to drinking in general, it should be clear that the prevention of alcoholism and chronic drinking problems should not be seen in isolation from the drinking culture at large. Hence, in light of the evidence, it is difficult to avoid the conclusion that if one wants to reduce the number of heavy drinkers and alcoholics in a country, then

one needs to pay attention to the general level of consumption in society.

Still, it is frequently argued that measures affecting mainly the 'normal' drinker are useless in preventive programs. This argument overlooks the fact that the 'normal' drinkers create the drinking culture in which the heavy drinkers are embedded. Ordinary drinking, so to speak, forms the cultural foundation of abuse, as is seen from the fact that population rates of abuse increase and decrease in direct response to the overall level of drinking in the culture.

If preventive efforts are exclusively targeted at heavy drinkers and potential heavy drinkers, with no attention given to 'normal drinking,' then there is the risk that any effects will rapidly become extinguished, since some of the mechanisms that caused the problem in the first place are left unchanged. A variant of this problem is well-known in alcoholism treatment — and the treatment of all substance abuse for that matter. Relapse is often related to direct or indirect social pressure (Marlatt and George, 1984), and returning to a wet environment significantly increases the risk of continuation of the drinking problem (Hörnqvist *et al.*, 1988).

A preventive policy that focuses exclusively on alcoholics would also easily fail to have an impact on intermediate drinkers — i.e. drinkers who run a certain risk for experiencing drinking problems, but who are not really alcoholics in the clinical sense. Since this group seems to be fairly large, intermediate drinkers may contribute substantially to aggregate problem rates.

This fact was demonstrated by Moore and Gerstein (1981, p. 42-47) on the basis of U.S. data. The bulk of problems were found among moderate drinkers, rather than among heavy drinkers. Obviously, the individual moderate drinkers have a much smaller risk than the individual heavy drinkers, but since there are many more moderate drinkers, the total number of problematic cases becomes larger in this group. Similar results have been reported from other countries as well (Skog, 1985d; Kreitman, 1986). Kreitman (1986) has called this effect the "prevention paradox." He discusses the implications of such findings for prevention, showing that even quite dramatic reductions in the prevalence of heavy drinking have very modest effects on problem rates, unless a parallel reduction occurs among moderate drinkers. It should be observed that this effect occurs mainly in connection with social and behavioral problems and problems of acute intoxication. As was pointed out earlier (section 2), the conclusion cannot be generalized to all alcohol-related problems. In fact, in the case of cirrhosis the result does not apply (cf. Pequignot *et al.*, 1978).

To sum up, there are two strong reasons why preventive efforts in relation to alcohol problems should not focus exclusively on alcoholism and alcohol dependence. Firstly, a large fraction of the

problem is to be found outside the ranks of the alcoholics, and secondly the prevalence of alcoholism is intimately connected to drinking in general.

As these facts have become acknowledged, the regulation of per capita alcohol consumption has come into focus in prevention. A direct manifestation of this is the fact that WHO-Euro (the World Health Organization's European office) has recommended to member states a 25 per cent reduction in per capita alcohol consumption by the year 2000, as part of the health-for-all program.

However, not all alcohol problems are due to chronic heavy drinking, and the per capita consumption level is not the only aggregate level variable of interest. Recognizing the relevance of per capita alcohol consumption to public health should not lead us to ignore other aspects of drinking.

As was noted earlier, the relationship between the general level of consumption and the prevalence of heavy drinking is not a law of nature. It is an approximate, empirical regularity, produced by socio-cultural mechanisms. Under certain circumstances, changes in per capita consumption may fail to be accompanied by parallel changes in rates of heavy drinking, and sometimes changes in heavy drinking and problem rates may occur without parallel changes in the general level of consumption. Hence, one should not commit the error of concluding that since per capita consumption is important, any measure that fails to affect per capita consumption is without effect.

The need to include dimensions other than per capita consumption level becomes even more obvious when we turn to problems of acute intoxication. Here, preventive efforts need to be targeted at both quantitative and qualitative dimensions of the drinking culture. Hence, it is hard to avoid the conclusion that in order to be 'complete', a sensible preventive strategy needs to include measures aimed both at the general level of consumption in society, at the drinking patterns in a more qualitative sense — in particular, acute intoxication and cultural rules for drunken comportment — and at the contexts and social situations in which people drink. These foci for primary prevention then need to be supplemented by measures focused on particular risk groups. Fortunately one does not have to choose one single item from this menu. A sensible combination of all elements would be optimal, provided that the strategy is based on an analysis of the nature of the problem in each particular country, and acknowledging the real potential of each preventive measure.

APPENDIX 1. AGGREGATE RELATIONS FOR DIFFERENT RISK FUNCTIONS

Let X denote annual consumption level, while R(X) denotes the risk experienced by a person with consumption level X. Furthermore, F(X) denotes the cumulative distribution function, i.e. the proportion of the population drinking less than X. The annual death rate D may then be written as an integral of the risk function with respect to the distribution function

$$D = \int_0^\infty R(X)\, dF$$

1. Linear risk function

Let

$$R(X) = a + bX$$

Here a and b are constants: a is the risk experienced by an abstainer, while b is the increase in risk produced by an increase in intake of one unit of alcohol. It immediately follows that

$$D = a + b\mu$$

where μ by definition is the mean consumption. Hence, when the individual risk increases linearly with consumption, the death rate in the population depends directly on the mean consumption level. The shape of the distribution is immaterial in this case. In fact, the mortality rate for two populations with the same mean consumption but very different distributions will be identical in this case.

2. Convex risk function

Let

$$R(X) = a + bX + cX^2$$

then

$$D = a + b\mu + c(\mu^2 + \sigma^2)$$

Hence, in this case the death rate depends on both the mean consumption level and the variance of the distribution (σ^2). Provided that the constants a, b, and c are all positive, this risk function is increasing and convex (curved upwards), resembling the curves reported for several alcohol-related diseases. Since the risk increases disproportionately quickly at high consumption levels, population mortality rate will depend on the prevalence of heavy drinkers. Hence, comparing two populations with the same mean, but different dispersions, the one with the largest variance will have the highest death rate.

If θ denotes the coefficient of variation (i.e. $\theta = \sigma/\mu$), we may write

$$D = a + b\mu + c(1 + \theta^2)\mu^2$$

Provided that θ remains unchanged while the mean consumption increases, the death rate in the population will increase as a quadratic function of μ. If θ tends to decrease slightly as μ increases, as available data suggest (Bruun *et al.*, 1975), the death rate will increase somewhat more slowly than the quadratic function.

3. U-shaped risk function

The quadratic risk function is U-shaped, of the type hypothesised for ischemic heart disease, when $b<0$ (cf. **Figure 1**). The minimum risk is attained when the individual consumption level is

$$X = -b/2c$$

Consider the case where θ remains unchanged when μ increases. This is the case when drinkers at all levels change their intake at the same relative rate, i.e. percentagewise. Under these circumstances the population mortality rate will be at its minimum level when the mean consumption level is

$$\mu = -b/[2c(1 + \theta^2)]$$

This population minimum level is less than the individual minimum level by a factor of $(1 + \theta^2)$. If the coefficient of variation should be $\theta = 2$, then optimum consumption level for the population becomes only one-fifth of the optimum for an individual.

At the level of individuals, the relative risk (RR) at the optimal consumption level becomes

$$RR_{min} = 1 - b^2/4ac$$

Hence, the risk is decreased by $100b^2/4ac$ per cent compared to an abstainer.

Provided that θ remains unchanged when average consumption changes, one finds that a population drinking at its optimal level would have the relative mortality (RD) rate (i.e. relative to a population of abstainers)

$$\mathbf{RD_{min} = 1 - b^2/[4ac(1 + \theta^2)]}$$

Hence, the population decrease in mortality is smaller than the individual decrease by a factor of $(1 + \theta^2)$. Unless one finds a way of reducing the dispersion of the distribution, or θ, the aggregate advantage may turn out to be fairly small.

APPENDIX 2. THE EFFECT ON AGGREGATE PROBLEM RATES OF CHANGES IN THE STABILITY OF DRINKING CAREERS

Let the risk for a certain chronic disease be a convex, quadratic function of cumulative consumption Y, e.g.

$$R(Y) = a + cY^2$$

The mortality rate then becomes

$$D = a + c(v^2 + m^2)$$

Here m denotes the mean and v^2 denotes the variance of the cumulative consumption measure.

Furthermore, let the cumulative consumption measure Y be a weighted sum of present and past consumption levels over a period of L years. We let X(0) denote present consumption, X(1) last year's consumption, and in general X(t) the consumption level t years ago. Lastly, we let w(t) denote the etiological weight attributed to the consumption level t years ago with respect to present risk. Without loss of generality we let the weights sum to unity. Then we may write

$$Y = w(0)X(0) + w(1)X(1) + ... + w(L)X(L)$$

Now consider for the sake of simplicity that the mean consumption level and the distribution of annual intake remains stable for a long period of time, and let μ and σ^2 denote the mean and the variance of the distribution of annual intake.

Under these circumstances the cumulative consumption measure Y has the same mean as annual intake, i.e.

$$m = \mu$$

However, the variance of cumulative consumption is not identical to the variance of annual intake. Rather, v^2 depends in a complicated way on how individual drinking varies over time, as measured by the correlations between intake in different years.

A simple example can illustrate this. Consider a (highly hypothetical) disease where only this year's and last year's consumption matter, and that the two years have the same

etiological importance. Hence $Y = [X(0) + X(1)]/2$ and the variance of Y becomes

$$v^2 = (1/2)(1 + r)\sigma^2$$

where r denotes the correlation between this year's and last year's consumption. Furthermore, the mortality rate becomes

$$D = a + c[1 + \theta^2(1 + r)/2]\mu^2$$

where θ denotes the coefficient of variation in the distribution of annual consumption.

From this, one observes that the variance and the mortality rate depend on the degree of correlation, in addition to depending on the mean and the dispersion of the distribution. The larger the correlation, the higher the mortality rate. The variance may be as small as $\sigma^2/2$ or as large as σ^2 in this case, i.e. it may differ by a factor of 2. If $\theta=1$, the alcohol-related mortality rate (i.e. D–a) will differ by a factor of 1.33 — that is, perfect correlation will give a 33 per cent higher mortality rate than a zero correlation. If $\theta=2$, the difference in mortality becomes greater, namely by a factor of 1.67.

When the time lag is much larger, the impact of the correlation also becomes much larger. We shall illustrate this by considering a case where the weights decline exponentially, i.e. $w(t)=\propto t$. This has been suggested as a first approximation for the risk function of liver cirrhosis (Skog, 1985c). Furthermore, we assume that the correlation between consumption levels some years apart decreases with increasing distance in time, namely as $r(t)=ßt$. (This implies that the changes in drinking are described by a first-order autoregressive process.)

Under these circumstances, the variance of the cumulative consumption measure can be shown to be

$$v^2 = [(1 - \propto)(1 + \propto ß)][(1 + \propto)(1 - \propto ß)]^{-1}\sigma^2$$

Table 3 shows how this variance increases with increasing correlations (ß), i.e. increasing regularity in drinking careers, for diseases with a fairly small ($\propto=.5$), a medium ($\propto=.8$), and a large ($\propto=.9$) time lag. The latter appears to be a reasonable approximation for cirrhosis of the liver (Skog, 1985c).

The effects on the variance of the cumulative consumption measure are quite dramatic, particularly when the time lag is large. For instance, if two populations have one-year correlations of .6 and .8 respectively, the latter will have an 84 per cent higher variance than the former (0.323 vs. 0.176). Provided that the risk function is quadratic, as presumed

Table 3. Variance of the cumulative consumption measure for different degrees of correlation between consumption in successive years and for different degrees of time lag in the cumulative measure. (Units of measurement have been chosen so as to make the variance of annual intake equal to one.)

	TIME LAG		
CORR.	SHORT	MEDIUM	LONG
.0	.333	.111	.052
.1	.368	.130	.063
.2	.407	.153	.075
.3	.450	.181	.091
.4	.500	.215	.111
.5	.555	.259	.138
.6	.619	.316	.176
.7	.692	.393	.231
.8	.777	.506	.323
.9	.878	.682	.501
1.0	1.000	1.000	1.000

above, the difference in alcohol-related mortality rate will be 13 per cent if $\theta=1$ and 35 per cent if $\theta=2$.

Clearly, changes or inter-cultural differences in the stability of drinking careers may have a quite significant effect on mortality rates for chronic alcohol-related diseases with convex risk functions. Even more strongly convex functions, such as the exponential function, may produce even larger effects.

REFERENCES

Bronetto, J. *Alcohol Price, Alcohol Consumption and Death by Liver Cirrhosis in the United Kingdom.* Mimeo. Toronto: Addiction Research Foundation 1961.

Bruun, K. *et al., Alcohol Control Policies in Public Health Perspective.* Helsinki: Finnish Foundation for Alcohol Studies 1975.

Cahalan, D. and Room, R. *Problem drinking among American men.* New Brunswick, N.J.: Rutgers Center of Alcohol Studies 1974.

Clark, W. Loss of control, heavy drinking, and drinking problems in a longitudinal setting. *Journal of Studies on Alcohol* 37: 1256-1290, 1976.

Clark, W. and Cahalan, D. Changes in problem drinking over a four year span. *Addictive Behaviors* 1: 251-259, 1976.

Cohen, M. *et al.* Moderate drinking by chronic alcoholic. *Journal of Nervous and Mental Disease* 153: 434-444, 1971.

Davies, D.L. Normal drinking in recovered alcohol addicts. *Quarterly Journal of Studies on Alcohol* 23: 94-104, 1962.

de Lint, J. and Schmidt, W. The distribution of alcohol consumption in Ontario. *Quarterly Journal of Studies on Alcohol* 29: 968-973, 1968.

Drew, L.R.H. Alcoholism as a self-limiting disease. *Quarterly Journal of Studies on Alcohol* 29: 956-967, 1968.

Duffy, J.C. The distribution of alcohol consumption — 30 years on. *British Journal of Addiction*, 81: 735-174, 1986.

Durkheim, E. *Suicide* New York: Free Press 1951.

Edwards, G. A later follow-up of a classical case series: D.L. Davies' 1962 report and its significance for the present. *Journal of Studies on Alcohol* 36: 181-190, 198.5

Fillmore, K. and Midanik, L. Chronicity of drinking problems among men: A longitudinal study. *Journal of Studies on Alcohol* 45: 228-236, 1984.

Frankel, B.G. and Whitehead, P.C. *Drinking and Damage: Theoretical Advances and Implications for Prevention.* New Brunswick: Rutgers Center of Alcohol Studies 1981.

Gove, W.R. & Hughes, M. Re-examining the ecological fallacy: A study in which aggregate data are critical in investigating the pathological effects of living alone. *Social Forces*, 58: 1157-1177, 1980.

Grant, B., Noble, J. and Malin, H. Decline in liver cirrhosis mortality and components of change: United States, 1973-1983. *Alcohol Health and Research World* no. 3, 10:66-69, 1986.

Hauge, R. and Irgens-Jensen, O. The relationship between alcohol consumption, alcohol intoxication and negative consequences of drinking in four Scandinavian countries. *British Journal of Addiction* 81: 513-524, 1986.

Heather, N. and Robertson, I. *Controlled Drinking*. London: Methuen 1983.

Herd, D. Migration, cultural transformation and the rise of black liver cirrhosis mortality. *British Journal of Addiction* 80: 397-410, 1985.

Hörnqvist, J.O., Hansson, B. and Åkerlind, I. The working capacity of the alcohol abuser. *Scandinavian Journal of Social Medicine* 16: 27-33, 1988.

Ipsen, J. An epidemic of infectious hepatitis, predominantly of adults, and highly fatal for elderly women. *American Journal of Hygiene* 51: 255-263, 1950.

Jellinek, E.M. *The Disease Concept of Alcoholism* New Haven: Hillhouse Press 1960.

Johansson, B.G. and Medhus, A. Increase in plasma alpha-lipoproteins in chronic male alcoholics during and after abuse. *Acta Medica Scandinavica* 195: 273-277, 1974.

Knibbe, R.A., Drop, M.J., Van Reek, J. and Saenger, G. The development of alcohol consumption in the Netherlands: 1958-1981. *British Journal of Addiction* 80: 411-419, 1985.

Knupfer, G. Drinking for health: the daily light drinker fiction, *British Journal of Addiction* 82: 547-555, 1987.

Kreitman, N. Alcohol dependence and the preventive paradox. *British Journal of Addiction* 81: 353-363, 1986.

Ledermann, S. *Alcool, Alcoolism, Alcoolisation.* Vol. I Paris: Presses Universitaires de France 1956.

Ledermann, S. *Alcool, Alcoolism, Alcoolisation.* Vol. II Paris: Presses Universitaires de France 1964.

Lelbach, W.K. Epidemiology of alcoholic liver disease. In: Popper,H. & Schaffner,F. (Eds.) *Progress in Liver Disease.* Vol. V. New York: Grune and Stratton, 1976 (pp. 494-515).

Lenke, L. Totalkonsumtionens betydelse för alkoholskadesituasjonen i Sverige. *Sociologisk Forskning* 20: 45-57, 1983.

MacAndrew, C. and Edgerton, R.B. *Drunken Comportment* Chicago: Aldine 1969.

Mäkelä, K. Levels of consumption and social consequences of drinking. In: Israel, Y. *et al.,* (eds.) *Research Advances in Alcohol and Drug Problems*, Vol. 4, New York: Plenum Press 1978 (pp. 303-348).

Mäkelä, K. *et al., Alcohol, Society and the State.* Toronto: Addiction Research Foundation 1981.

Mäkelä, K. and Mustonen, H. *Positive and negative experiences related to drinking as a function of annual alcohol intake.* Paper presented at the Epidemiology Section of the International Council of Addiction and Alcoholism, Aix-en-Provence, June 1987.

Mann, R.E., Smart, R., Anglin, L. and Rush, B.R. Are decreases in liver cirrhosis rates a result of increased treatment for alcoholism? *British Journal of Addiction*, 83: 683-688, 1988.

Marlatt, G.A. and Nathan, P.E. *Behavioral Approaches to Alcoholism.* New Brunswick, N.J.: Rutgers Center of Alcohol Studies 1978.

Marlatt, G.A. and George, W.H. Relapse prevention: Introduction and overview of the model. *British Journal of Addiction* 79: 261-273, 1984.

Miller, P.M. *Behavioral Treatment of Alcoholism.* New York: Pergamon Press 1976.

Moore, M.H. and Gerstein, D.R. *Alcohol and Public Policy: Beyond the Shadow of Prohibition.* Washington, D.C.: National Academy Press 1981.

Norström, T. The impact of per capita consumption on Swedish cirrhosis mortality. *British Journal of Addiction*, 82: 67-75, 1987a.

Norström, T. The abolition of the Swedish alcohol rationing system: Effects on consumption distribution and cirrhosis mortality. *British Journal of Addiction*, 82: 633-642, 1987b.

Petersen, J. Skinhøj, P. and Thorsen, T. An epidemic of cirrhosis in Danish women revisited. *Scandinavian Journal of Social Medicine*, 14: 171-178, 1986.

Pequignot, G., Tuyns, A.J. and Berta, J.L. Ascitic cirrhosis in relation to alcohol consumption. *International Journal of Epidemiology* 7: 113-120, 1978.

Poikolainen, K. *Alcohol Poisoning Mortality in Four Nordic Countries.* Helsinki: Finnish Foundation for Alcohol Studies 1977.

Polich, J.M., Armor, D.J. and Braiker, H.B. *The Course of Alcoholism. Four Years After Treatment.* New York: Wiley 1981.

Popham, R.E. Indirect methods of alcoholism prevalence estimation: A critical evaluation. In: Popham, R.E. (ed.) *Alcohol and Alcoholism*, Toronto: University of Toronto Press 1970.

Popham, R.E., Schmidt, W. and Israel, Y. Variation in mortality from ischemic heart disease in relation to alcohol and milk consumption. *Medical Hypotheses* 12: 321-329, 1983.

Romelsjö, A. Decline in alcohol-related inpatient care and mortality in Stockholm county. *British Journal of Addiction*, 82: 653-663, 1987.

Room, R. Measurement and distribution of drinking patterns and problems in general populations. In: Edwards, G., Gross, M.M., Keller, M., Moser, J. and Room, R. *Alcohol-related Disabilities.* Geneva: World Health Organization 1977 (pp. 61-87).

Roizon, R. *et al.* "Spontaneous remission" among untreated problem drinkers. In: Kandel, D.B. (ed) *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* New York: John Wiley 1978 (pp. 197-221).

Sarles, H. Alcohol and the pancreas. In: Tygstrup, N. & Olsson, R. (Eds.) Alcohol and Disease. *Acta Medica Scandinavica Symposium Series No. 1.* Stockholm: Almqvist & Wiksell International, 1985 (pp. 235-250).

Schmidt, W. and Bronetto, J. Death from liver cirrhosis and specific beverage alcohol consumption: an ecological study. *American Journal of Public Health* 52: 1473-82, 1962.

Schmidt, W. Agreement, disagreement in experimental, clinical and epidemiological evidence on the etiology of alcoholic liver cirrhosis: a comment. In: Khanna, J.M., Israel, Y. and Kalant, H. (eds.) *Alcoholic liver pathology*, Toronto: Addiction Research Foundation 1975 (pp. 19-30).

Schmidt, W. and Popham, R.E. The single distribution theory of alcohol consumption; a rejoinder to the critique of Parker and Harman. *Journal of Studies on Alcohol* 39: 400-419, 1978.

Skog, O.-J. *Modeller for drikkeatferd.* (Models of drinking behavior) SIFA-mimeo no. 32. Oslo: National Institute for Alcohol Research 1979.

Skog, O.-J. Social interaction and the distribution of alcohol consumption. *Journal of Drug Issues* 10: 71-92, 1980a.

Skog, O.-J. Liver cirrhosis epidemiology: Some methodological problems. *British Journal of Addiction* 75: 227-243, 1980b.

Skog, O.-J. *The distribution of alcohol consumption. Part I. A critical discussion of the Ledermann model.* SIFA-mimeo no. 64. Oslo: National Institute for Alcohol Research 1982.

Skog, O.-J. The collectivity of drinking cultures. A theory of the distribution of alcohol consumption. *British Journal of Addiction* 80: 83-99, 1985a.

Skog, O.-J. The wetness of drinking cultures: A key variable in epidemiology of alcoholic liver cirrhosis. *Acta Medica Scandinavica Supplement* 703, 157-184, 1985b.

Skog, O.-J. The risk function for liver cirrhosis from lifetime alcohol consumption. *Journal of Studies on Alcohol* 45: 199-208, 1985c.

Skog, O.-J. Hva bestemmer omfanget av alkoholskader? (What determines the volume of alcohol problems?) In: Arner, O.; Hauge, R. and Skog, O.-J. (eds.) *Alkohol i Norge* (Alcohol in Norway) Oslo: Universitetsforlaget 1985d.

Skog, O.-J. *The distribution of alcohol consumption. Part III. Evidence for a collective drinking culture.* SIFA-mimeo no. 8/85. Oslo: National Institute for Alcohol Research 1985e.

Skog, O.-J. Trends in alcohol consumption and violent deaths. *British Journal of Addiction* 81: 365-379, 1986a.

Skog, O.-J. The long waves of alcohol consumption: a social network perspective on cultural change. *Social Networks* 8: 1-32, 1986b.

Skog, O.-J. Registrert drukkenskap i Norge 1931-1985 (Arrests for drunkenness in Norway 1931-1985). *Alkoholpolitik* 4: 1-9, 1987.

Skog, O.-J. Alcohol and suicide - Durkheim revisited. *Acta Sociologica* 34: 193-206, 1991.

Smart, R.G. Recent international reductions and increases in liver cirrhosis deaths. *Alcoholism - Clinical and Experimental Research* 12: 239-242, 1988.

Smart, R.G. and Mann, R.E. Large decreases in alcohol-related problems following a slight reduction in alcohol consumption in Ontario 1975-83. *British Journal of Addiction* 82: 285-291, 1987.

Tuyns, A.J., Pequignot, G. and Jenssen, O.M. Le cancer de l'esophage en Ille et Vilaine en fonction des niveaux de consommation d'alcool et de tabac. Des risques qui se multiplient. *Bull. Cancer* 64: 46-60, 1977.

Ullman, A.D. Sociocultural backgrounds of alcoholism. *Annals of the American Academy of Political and Social Science*. 315: 48-54, 1958.

Vaillant, G.E. *The Natural History of Alcoholism: Causes, Patterns and Paths to Recovery*. Cambridge, Mass.: Harvard University Press 1983.

Whitehead, P.C. and Harvey, C. Explaining alcoholism; an empirical test and reformulation. *Journal of Health and Social Behavior* 15: 57-65, 1974.

Williams, P., Robinson, D. and Balley, A. High-density lipoprotein and coronary risk factors in normal men. *Lancet* I: 72-75, 1979.

Yano, K., Rhoads, G.G. and Kagan, A. Coffee, alcohol and coronary heart disease among Japanese men living in Hawaii. *New England Journal of Medicine* 297: 405-409, 1977.

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

Dependence: Uncertainties, Mysteries and Doubts

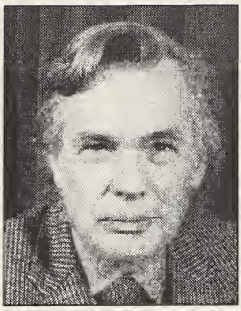
GRIFFITH EDWARDS

ADDICTION RESEARCH UNIT

NATIONAL ADDICTION CENTRE

INSTITUTE OF PSYCHIATRY

LONDON, ENGLAND



GRIFFITH EDWARDS

A graduate of Oxford University, Griffith Edwards was introduced to the clinical study of alcoholism as a psychiatric resident at the Maudsley Hospital. Dr. Edwards has written and edited a number of works on alcohol and drug issues, including *Pharmacological Treatments for Alcoholism, The Nature of Drug Dependence* and *Unreason in an Age of Reason*. He is editor of the prestigious *British Journal of Addiction*, Professor of Addiction Behavior at the University of London, Director of the Addiction Research Unit of the National Addiction Centre and course director for the NAC's international Diploma in Addiction Behavior. Dr. Edwards won the 1979 Jellinek Award for alcohol studies, along with D.L. Davies, and in 1987 was awarded the CBE for services to social science and medicine.

INTRODUCTION

Five types of question are posed around withdrawal symptoms: 1. The reasons for the fluctuations over time in medical awareness of alcohol's capability to produce withdrawal symptoms; 2. What type of drinking schedules instigate withdrawal experience? 3. Do withdrawal phenomena get worse over time? 4. The 'addiction memory' question; and lastly 5. Is withdrawal a reinforcer for drinking behavior? These issues point to a need for a renascence in the kind of research that is based on clinical observation and on listening to what patients have to tell, with such lines of work brought into much closer contact than before with corresponding laboratory investigations.

'NEGATIVE CAPABILITY'

The field of alcohol studies has at times been disadvantaged by a too hasty commitment to fixed positions. We would do well more often to heed the advice of John Keats (1817) as to the value which attaches to the suspension of quick judgement. "Negative capability," he said, "is when a man is capable of being in uncertainties, mysteries, doubts, without any irritable reaching after fact and reason...."

Let us then try out our own negative capabilities and select as substance for such testing a specific topic — the issue of alcohol withdrawal. Here by way of introduction to a journey around that difficult territory are two statements selected for the tension that they set up between them. The first quotation is from Fingarette's *Heavy Drinking: The Myth of Alcoholism as a Disease* (Fingarette, 1988).

> "But although morning drinking is a meaningful clue about a drinker's way of life, to call it a symptom imputes a disease and connates a medical phenomenon. By way of analogy: If we heard a person frequently discussing business over breakfast, we would not call this activity a 'symptom' of an ambitious commitment to business. But we might consider this breach of cultural norms regarding polite breakfast conversation to be a reliable sign of some significant dimension of the drinker's way of life."

A couple of pages later Fingarette continues:

> "Nor is there some mysterious demon afoot — at each point the heavy drinker's conduct is an intelligible version of normal

cognitive and emotional behavior."

Thus, for that particular author there is no mystery attaching to the morning drink, and the matter is disposed of by the assertion that such behavior is to be conceived and explained as being on a level with mixing the stock market with toast and marmalade.

Here is the second introductory statement and it's from a very different authority:

> "Before long I was back on my full amount again... same old sweats and nausea. Now up till then I never drank in the mornings. Even if I went to sleep and left a glass of whisky, which I often did, I'd always pour it back in the bottle the next morning. Didn't want to touch it. This used to puzzle me — why don't I drink in the morning, and a few of the old timers used to say, it'll come. Well, some time about July or August, I found that I just couldn't get out of bed. When I put my feet down they were like jelly. And I thought, well, this swig might help me on my way and I took a swig and was sick a bit and after a few minutes, lovely, back on your legs. It shows you how naive I was at that time, yet I had enough experience of drink. I thought that swig would last me till lunchtime. Well, I'd been up about half-an-hour and I wanted another one... and when I got to work at 8 a.m., I wanted another drink. The place where I got it from didn't open till 9 a.m. Jesus, I'll always remember this, I had to wait till 9 a.m., sweating. That's when my morning drinking started."

This latter quotation comes from a series of tape-recorded research interviews conducted by the Addiction Research Unit. The subject was a 50-year-old man who had suffered from alcohol dependence in severe degree. To introduce into scientific debate the unvalidated retrospection of one selected drinker might be viewed as unusual or even out of order. It could on the contrary be argued that the scientist should be willing to return again and again to exactly such material.

Pause for a moment and contemplate some of the questions to which that man might be seen as pointing. In the process we can identify five headings which will provide a useful framework for the remainder of this paper.

1. What in historical terms has science made of stories of the kind that this man now relates? The evidence points in fact to an ebb and flow over the decades in expert awareness, a 'now you see it, now you don't' state of affairs.

2. What type of drinking schedules or other influences determine the first onset of the withdrawal experience and withdrawal relief drinking? Why did that man 'always pour it back into the bottle' until that one particular morning?

3. Do withdrawal phenomena get worse over time and if so, why, and at what pace?

4. Is the individual who has once experienced severe withdrawal symptoms more prone to develop them again in a subsequent drinking bout than the person who has not previously manifested this behavior — the 'addiction memory' question?

5. Is withdrawal a reinforcer for drinking behavior?

Before tackling each of these headings in turn, it is necessary to clarify the boundaries of what will be covered. For the sake of brevity the main focus will be on withdrawal rather than on the closely related issue of tolerance. In many places where the word 'withdrawal' is used, it might indeed be more proper to employ the phrase 'tolerance and withdrawal.' The discussion will furthermore be limited to alcohol and alcohol withdrawal, although the same themes could (with variation) be analysed in relation to many other types of substance use.

LESSONS OF HISTORY

Today, the question of whether alcohol is a drug that can give rise to tolerance and withdrawal — to neuroadaptation — is one that is likely to be deemed redundant by most people. The universal answer from the standard textbooks would be that the alcohol-dependent human subject or experimental animal will, on cessation of sufficiently intensive exposure to alcohol, experience withdrawal symptoms as evidenced by a spectrum of manifestations extending all the way from mild or severe tremors, sweats, nausea and affective disturbance, through to delirium tremens or grand mal seizures. Look, however, along the library shelves and it becomes evident that in the very recent past the received wisdom was exactly the opposite of what we accept today.

Eras of clinical awareness

The full history of the way in which medicine and science have regarded alcohol withdrawal symptoms has yet to be written and for present purposes the chronicling of events before the 1940s will be restricted to identification of a few important milestones. Hippocrates noted the existence of alcohol withdrawal phenomena. Leaping foward a mere couple of thousands years or so one finds Thomas Trotter (1804) writing thus in his famous *Essay on Drunkenness*:

◆

> "The head and hands of some inebriates, particularly in the morning, shake and tremble; but regain their usual strength, and become steady, as the dose of stimulant is repeated."

That passage reveals an unequivocal identification not only of withdrawal tremors but also withdrawal relief drinking. Trotter did not however favor the withdrawal explanation of delirium tremens.

In the nineteenth and earlier part of the twentieth century, withdrawal symptoms continued to be recognized as a clinical reality — for instance, Magnus Huss made reference to them in his *Alcoholismus Chronicus* (Huss, 1849). Somewhat later a detailed discussion of this topic is to be found in a paper written by Francis Hare, Medical Superintendent of the Norwood Sanitarium, which was published in 1910 in the *British Journal of Inebriety*, the earlier title of the *British Journal of Addiction*. He wrote:

> "It may be stated generally that just in proportion to the degree of tolerance of alcohol which has been acquired, so is the danger of sudden withdrawal. Acquired tolerance implies intolerance of sudden or very rapid withdrawal. The chief complications which are liable to arise are delirium tremens and alcoholic epilepsy, or both. Such never occur except under too rapid withdrawal...."

The MRC: Awareness with provisos

At some point in the first half of the 20th century, clinical awareness of withdrawal phenomena faded, and Wallgren & Barry (1970) go so far as to state that the concept of a withdrawal illness 'gradually fell into oblivion.' As has often been the situation in relation to other clinical issues, a fuller analysis of the historical material might well show that such a judgement is too general, but an inspection of a few contemporary texts certainly suggests that over several decades the dependence potential of alcohol was generally downplayed. For example, the Medical Research Council's influential report on *Alcohol: its Action on the Human Organism* (MRC, 1938), first published in 1918, stated as follows:

> "... the continued use of alcohol produces tolerance though never to the extent that may be induced in the morphine habit. And alcohol also may induce such a condition in the drinker that its complete and sudden withdrawal will give rise to a feeling of urgent need or 'craving' for the drug, which is usually regarded as, and probably is, at least in part, a genuine abstinence symptom. Though very lurid accounts of this craving are sometimes given by neurotic drinkers, it would appear that ordinarily the feeling, insofar as it

> really is such a symptom, is very much weaker than the craving which habitual users of other drugs experience as a result of abstinence."

Thus the MRC's report made mention of the abstinence syndrome, but there was a paucity of clinical detail and their appraisal of the importance of those symptoms was hedged around with provisos. There was a short discussion of delirium tremens under the heading 'Delirium tremens not due to abstinence.' The committee from which this report emanated included such authoritative figures as Henry Dale, William McDougall, Frederick Mott and Charles Sherrington.

Yale school: Awareness lost

Jellinek's *Alcohol Addiction and Chronic Alcoholism*, edited on behalf of the Scientific Committee of the Research Council on Problems of Alcohol and published in 1942, provides another landmark (Jellinek, 1942). In an otherwise remarkably detailed and comprehensive series of reviews, what stands out is the lack of interest in withdrawal phenomenology other than in relation to whether patients should be abruptly withdrawn from alcohol or 'tapered off,' together with the related question as to whether fits or DTs could be caused by withdrawal. In a chapter in this book Bowman & Jellinek (1942) stated boldly that "alcoholic epilepsy" was not an "alcoholic mental disorder," but due to some other underlying cause. The view that withdrawal was the cause of delirium tremens had, they reported, been "virtually discarded" in the U.S., although "in Europe, particularly in France, it still has some unquestioning adherents."

Alcohol, Science and Society, a collection of lectures given at the Yale Summer School, was first published in 1945 (Yale University Center of Alcohol Studies, 1945). The transcripts of the question-and-answer sessions that followed the formal lectures give lively insights into contemporary thinking. Here, for instance, was Howard H. Haggard (Haggard, 1945) dealing with a question on tolerance:

> "Very little is known about tolerance. There is no true toxicological tolerance to alcohol as there is for some alkaloids, such as nicotine and morphine."

And here was Haggard in a later session giving the views on withdrawal symptoms:

> Question: "Is there anything, physiologically speaking, to the idea that a person who has been on a severe alcoholic binge should taper off?"
>
> Haggard: "You perhaps have in

> mind the devastating effects and suffering that occur in the man addicted to morphine or other narcotics on abrupt withdrawal. Symptoms of such severity do not develop in the alcoholic. He may be miserable but he has not developed the altered physical state in which the cells of the body fail to function normally except in the presence of the drug."

Joliffe (1945) disposed of alcoholic epilepsy by invoking edema of the brain and did not even touch upon the possibility that the fits might be caused by withdrawal. A fair assessment of the views of these founding fathers of the post-war alcoholism movement might thus be that tolerance and withdrawal were matters only very peripheral to the view of alcoholism that they were at that time developing. Indeed, they were moving somewhat backwards even compared with the rather timid position taken in the MRC's report.

Awareness reborn: Two crucial factors

The turning point in awareness came in the 1950s with the publication of two crucial studies. Victor and Adams (1953) provided data that described the time-course and clinical spectrum of alcohol withdrawal. These data did much to establish that convulsions and delirium tremens could be brought about by withdrawal and they gave precision to what had been the received opinion among Dr. Hare and his colleagues in an earlier generation. Just two years after the Victor and Adams paper, Isbell and his colleagues (1955) seemingly clinched the matter with their experimental study of the etiology of 'rum fits' and delirium tremens, conducted at Lexington. In the introduction to their paper they stated: "In the United States it is generally accepted that withdrawal of alcohol plays no role in the precipitation of delirium tremens."

Follow-through to 'The Disease Concept of Alcoholism'

The new American research evidence was to have a quick follow-through in terms of impact on discussions that were in progress at the World Health Organization in Geneva. WHO had issued reports in 1951 (WHO, 1951), 1952 (WHO, 1952) and 1954 (WHO, 1954), and was gradually moving toward a firmer position on the dependence potential of alcohol. The 1954 Expert Committee had for instance dealt with the matter thus:

> "... compulsive craving and dependence can develop in those individuals whose make-up leads them to seek and find an escape in alcohol. With this substance the

> personal make-up is the determining factor, but the pharmacological action plays a significant role."

Under the influence of the recent American research the 1955 report (WHO, 1955) then went significantly further:

> " ...though many of the events observed in alcoholism are parallel to many of the phenomena observed in opiate addiction, many important differences exist. The Committee feels, however, that recent evidence makes it appear that there is more resemblance between the responses to the withdrawal of alcohol and of opiates than was previously realized... following discontinuation of alcohol after a prolonged period of very heavy drinking, severe withdrawal symptoms, which in a limited proportion of cases include convulsions or delirium, or both, may occur.... Even though a marked degree of physical dependence on alcohol, as manifested by withdrawal symptoms, can develop, it occurs only after a prolonged period of very heavy drinking...."

This WHO statement constituted, as it were, the official *imprimatur* for the reality of alcohol withdrawal states. The tendency was still to take opiate dependence as the template and view alcohol dependence as being in some ways more minor, and the emphasis was still on gross alcohol withdrawal states rather than the everyday occurrence of more undramatic symptoms such as morning shakes. Nonetheless here was the categorical statement that 'serious symptoms' could result from alcohol withdrawal.

Jellinek was intimately involved in these WHO deliberations and the *Disease Concept of Alcoholism*, published in 1960, refers repeatedly to the series of WHO reports. This book also noted the significance of Victor and Adams (1953) and quoted extensively from Isbell *et al.* (1955). The thinking that Jellinek put forward in 1960 (Jellinek, 1960) constituted a radical shift from that of the 1940s and now, instead of withdrawal being of peripheral concern, it had become central to Jellinek's differentiation between his 'species' of alcoholism and a crucial plank in the formulation of the disease concept. Having ruled out alpha and beta alcoholism as candidates for the disease epithet, Jellinek continued as follows (p. 40):

> "This leaves us with the gamma, delta and epsilon species of alcoholism. The first two of these may come into consideration as diseases, since it is the adaptation of cell metabolism, acquired

> increased tissue tolerance and the withdrawal symptoms that bring about 'craving' and loss of control or inability to abstain...."

The significance given to withdrawal symptoms in this book is thus in very sharp contrast to the earlier stance of the Yale school. Its assertion that withdrawal symptoms *caused* 'craving' was not substantiated by evidence.

How is the ebb and flow to be explained?

Without making harsh judgements from the privileged position of hindsight, how is the ability of experts to forget or deny what was known earlier to many shrewd practitioners to be explained? One might speculate on the possible impact of changes in the relationship between the world of clinical observation and the high ground of scientific authority. Perhaps also the cultural-political background in some subtle fashion always makes science more or less able to see the facts before it. The evidence that over a period of time the best scientific minds were able to dismiss from their field of vision the facts that any shaking patient would have demonstrated or reported to them at any time, is perhaps some justification for earlier having introduced a patient's transcript into a review of the literature. 'Clinical experience' is by itself a notoriously unreliable arbiter on complicated issues, but to fail to test our received scientific wisdom against what our patients can tell us if only we will listen, may lead to at least equally serious wrong turnings.

WITHDRAWAL SYMPTOMS: HOW TO GET THEM

There is little or no research bearing on the question of how patients describe the initial onset of their withdrawal symptoms. Clinical impression would probably suggest a wide variation between subjects. Some may remember a fairly abrupt transition between a phase in which, despite long-continued heavy drinking, they only occasionally experienced withdrawal, and the ensuing phase in which morning shakes and nausea became an unpleasant, alien, everyday experience. Others will however picture the evolution in terms of much less clear-cut transitions. There is an issue as to the extent to which morning shakes are a phenomenon requiring a different kind of explanation than a social drinker's rebound symptoms, which can occur after a single drinking exposure. It will be assumed that the rebound type of happening needs to be examined largely in terms of only the immediately previous drinking experience, rather than in relation to the prior drinking behavior over more protracted time, and this discussion will focus on classical withdrawal.

There is a great deal of animal work that bears on the type of schedules that will produce withdrawal symptoms, and a

number of useful reviews of the animal literature have been published (e.g. Kalant *et al.*, 1971; Kalant, 1973; Cappell & LeBlanc, 1979; Cicero, 1980). The best interpretation of the evidence might be that in animals, tolerance or withdrawal symptoms can be produced within a few hours of first exposure (Goldstein, 1972, 1974), although the severity of the withdrawal state will be dose-related. In experimental animals there is thus no great mystery about the type of alcohol exposure schedule that will speedily and predictably induce withdrawal on cessation of drinking.

When human subjects were exposed to fixed-interval and high-dosage administration of alcohol in the experiment conducted by Isbell *et al.* (1955), severe withdrawal symptoms were induced on cessation after 48-87 days of alcohol exposure; but it is not easy to tell from the original account when during this period tremor first occurred. The heavy degree to which these patients were dosed with alcohol may of course have masked symptoms until the supply was abruptly curtailed at the end of the experiment. It should also be noted that all 10 experimental subjects were former morphine addicts and four of them had had previous drinking problems. The bearing of the body of work reported by such authors as Mendelson (1969), Mello (1972), Nathan *et al.* (1970) and Gross *et al.* (1973; 1975a,b) on the issue under discussion is limited by the fact that none of these researchers was administering alcohol to previously naive or moderately drinking subjects, but to volunteers with a history of severe previous alcohol dependence. Such work can reveal something about the possible time-course for reinstatement of withdrawal symptoms, but tells us nothing about how to get them in the first instance.

If the animal work has any relevance to the human condition one must however conclude that the experimentally feasible time-course for induction of withdrawal symptoms in a human subject would probably involve a far shorter period of exposure than the usual gap between first onset of heavy drinking (however defined) and appearance of established withdrawal symptoms as reported by most patients. In the former instance, one is often talking about a matter of only a few hours or days of exposure, while in the latter case one is perhaps usually witnessing a history of five to 15 years of heavy drinking before the onset of marked withdrawal.

What makes the watershed?

What factors might contribute to an explanation of why the typical patient may require a decade or more of heavy drinking to instate significant withdrawal experience when he probably has the biological capacity to bring about such experience in a fraction of that time? Here are some conjectures:

1. The usual dosage exposure or blood alcohol level attained in the natural setting may be subcritical for induction of

withdrawal symptoms. This might in part be due to the aversive consequences of very heavy drinking. Hence in terms of the dose-by-time formula the time base may have to be considerably extended to achieve the needed total exposure.

2. The pattern of exposure may be subcritical. The natural human drinking schedule is seldom an everyday affair with round-the-clock timetabling of the kind which Isbell *et al.* set up. This issue has been discussed by Turner (1980).

3. In the watershed period, an intensification in the alcohol exposure may occur either in terms of total intake or pattern of intake. These changes might then be expected to induce frank withdrawal where the scheduling was previously such as to be approaching some critical threshold. Increased spending power, or a cheapening in the real price of alcohol, might be important economic determinants and there is evidence that price will affect consumption even among heavy drinkers (Kendell *et al.*, 1983). Situational influences might include changes in the occupational, leisure or domestic patterning of drinking behavior.

4. There might be an alteration in the daily peak or the shape of the individual's blood alcohol curve due to such influences as change in diet or in absorption, despite the absence of any significant change in consumption. A partial gastrectomy might for instance alter the absorption curve quite radically. It is tempting to speculate that a decrease in liver capacity or metabolic tolerance might be influential at the margins of some withdrawal-induction threshold, although in a different context Kalant *et al.* (1971) have pointed out that clearance rates do not have a very large impact on the peak level attained when alcohol doses are in the heavier range.

5. The possible influence of changes in brain sensitivity should also be considered. Included here would be such factors as the aging process, brain damage, or the possible influence of other drugs including exposure to benzodiazepines.

Whatever the explanation for the onset of withdrawal symptoms, there must then also be separate questions relating to the onset of withdrawal-relief drinking (Orford, 1985). To what extent is the initiation of this behavior dictated by the sheer intensity of withdrawal? What is the significance for the natural history of dependence which may result from different cultures proposing very different norms on the propriety of morning drinking? Do some patients break through the barriers of convention and reach for the first morning drink earlier in the history of withdrawal experience than do others, and is this in part related to personality? Is the strategy of taking alcohol to relieve morning withdrawal usually hit upon by accident or is there social or imitative learning?

WITHDRAWAL AND ITS POTENTIAL FOR PROGRESSION

Do withdrawal symptoms get worse over time? The issue as to whether withdrawal is more easily *reinstated* with repeated cycles of drinking and abstinence will be left to a later section of this paper. In reality the two questions are linked.

The research background

So far as the present question is concerned, the light that can be thrown on it by animal experiments is valuable but probably limited. Such work is useful because it clearly demonstrates that in the short term the severity of the observable withdrawal syndrome increases to an asymptote over a period of perhaps, in some circumstances, only a few days (Goldstein & Arnold, 1976). The limitations in the inferences that can be drawn from these data derive partly from the obvious fact that the experiments do indeed involve a length of exposure to alcohol that is of an almost trivial order compared to the long years of the human subject's usual drinking career. The restriction set by the relatively short experimental period has often resulted in the investigator employing heavy doses of alcohol so as to maximize the likelihood of producing grossly evident withdrawal symptoms, including perhaps convulsions. It should also be remembered that this model can only offer a replication of certain objective signs but can tell us nothing about the progression of subjective distress. Another important limitation of the animal model includes difficulty to date in catching or manipulating within it the consequences of learning and conditioning on the march of withdrawal intensity.

The human experimental ward setting is similarly limited in that the data generally speak only to progression over at most two to three months of drinking (Isbell *et al.*, 1955), and deal more with reinstatement than with basic questions of progression. As for clinical investigations based on questionnaires and patient reports, there appears to be very little work of that kind bearing directly on the issue. Investigators may have been intimidated by the difficulties in retrospection, and no one has set up the formal prospective design.

In sum, neither animal nor human research can as yet do much to illuminate these seemingly rather basic questions as to how and in what circumstances withdrawal symptoms will show a march in their severity, the likely pace of that march, and the nature of any asymptote. And yet if one were to walk over from laboratory or library and sit down to debate the problem with a group of drinkers who had personal experience of alcohol dependence, one may conjecture that they would find this question pertinent and substantial.

Dissecting out the questions

Here is an attempt to dissect out in more detail some of the issues that, on closer inspection, might be found to cluster around the overall question of progression:

1. What is the likely sequence in the coming into play of individual withdrawal phenomena? Some patients seem, for instance, to report that morning nausea or vomiting considerably preceded the first experience of morning tremor, while for others the opposite occurs. For some the affective symptoms preceded the physical signs.

2. Once they have begun to occur, what are the possible patterns of progression for individual withdrawal phenomena or for any global rating (Friedman, 1980)? Progression must be considered not only in terms of *intensity* of an experience, but also as judged by *frequency* of occurrence. On the other hand it is possible that frequency and intensity are intimately related. Some patients appear to experience infrequent and low-level phenomena continuing over many years without significant progression and perhaps even with eventual *regression* in severity. It would therefore be a mistake to envisage the march as inexorable, and a model based exclusively on the idea of a one-way movement or 'phases of alcoholism' (Jellinek, 1952) is unlikely to contain all the observable happenings.

3. What factors may lead to the eventual occurrence of delirium tremens or withdrawal fits?

'ADDICTION MEMORY': A MYSTERY OR A MYTH?

At Alcoholics Anonymous meetings, the story of the long-time sober alcoholic who 'slips' is frequently heard. After a few days, he or she is allegedly back to experiencing withdrawal symptoms in full force, and is again engaging in desperately urgent withdrawal relief drinking. In terms of the dependence syndrome concept, dependence has been 'reinstated' (Edwards & Gross, 1976).

To many experienced clinicians that kind of story will illustrate one of the most evident but puzzling features of pathological drinking. In the first instance it will have taken the drinker many years of heavy alcohol exposure to instate severe withdrawal symptoms and yet after a long period of abstinence that patient may need only a few days of drinking to bring back his withdrawal symptoms at an intensity equal to anything he ever previously experienced.

If the classic story portrayed in this type of sketch accurately describes the nature of at least one type of slip, then we are indeed standing in the presence of an intriguing mystery. But before becoming too enthusiastic about the challenge posed by this question, it would be prudent to

pause and ask whether we are, in fact, in the presence only of a myth (Falk, 1983). How real is this alleged type of happening, or is it a story to be thrown out along with the rest of the 'one drink, one drunk' paraphernalia? Perhaps what is being seen here is often no more than the 'abstinence violation effect,' and something to be viewed as a self-fulfilling prophecy rather than a biological mystery (Marlatt, 1985).

The balanced judgement might be that we are probably in the presence of a mixture constituted in rather uncertain proportions of solid factual happenings that need to be explained and mythological overlay that needs to be dispelled. Learning effects have, however, a reality of their own and should not be dismissed as mythological (that was not Marlatt's intention). All that will be assumed for the present is that there is sufficient factual basis to make further scrutiny worthwhile.

'Addiction memory'

The term 'addiction memory' seems first to have been applied to the matters we are now discussing by Mello (1972), who wrote that "The time course of development and duration of such an agent-induced 'addiction memory' are not known." Kalant (1973) used the phrase 'carry-over' in a rather similar fashion and it is clear that he had in mind both tolerance and withdrawal. Edwards & Gross (1976) employed the term 'reinstatement' to describe the re-institution not just of tolerance and withdrawal but of the holistic clinical dependence syndrome. We will proceed by briefly examining first the animal and then the human research that bears on this question.

The animal work has received considerable attention in the various reviews to which reference has already been made in an earlier section of this paper. Within the usual experimental time-base the alcohol-experienced animal will acquire tolerance and manifest withdrawal more quickly than the alcohol-naive animal (Branchey *et al.*, 1971, Ellis & Pick, 1971; Kalant *et al.*, 1971; Walker & Zornetzer, 1974; Deutsch & Walton, 1977; LeBlanc & Cappell, 1975; Kalant *et al.*, 1978; Poldrugo & Snead, 1984). With repeated cycles over a few weeks a plateau is reached. Contrary results were obtained by Goldstein (1974), but the schedule she employed differed from that used by investigators who have obtained positive results (Baker & Cannon, 1979). The short time-base of most of this animal work inevitably again implies that statements on the duration of any carry-over effects may not have much relevance to the human condition, but Kalant *et al.* (1971) reported that with rats, "On the cessation of alcohol treatment, tolerance was found to revert gradually to normal during a period of two to three weeks."

Turning to what may be learnt from research conducted with patients, it may firstly be interesting to note an observation recorded by Mendelson and his colleagues. Mendelson & LaDou (1964a) described

the methodology of an experiment in which 10 male alcoholics were subjected to 24 days of fixed-interval drinking with bourbon given every four hours. There was a graded increase in dosage leading to the maximum daily intake of 40 oz. In one of the papers deriving from this project (Mendelson et al., 1964) the investigators described the changes in drinking behavior over time in the following terms, starting with the early phase when subjects were drinking 24 oz. daily:

> " ...by the fourth day... all the men stated that they felt no specific craving for alcohol, and that if the experiment were stopped, they would not experience any compulsion to continue drinking."

In the middle phase of the experiment, when subjects were consuming up to 30 oz. daily, the investigators observed that:

> "Associated with the heightened anxiety was a tendency toward more compulsive drinking behavior. The subjects drank their beverage faster than before, and occasionally requested more whisky before the usual time of administration. At the end of the period most of the subjects reported that they felt a definite need to continue drinking, although they denied the very strong feeling of need for alcohol that they experienced when drinking in real life situations."

Observations continued into the third phase when subjects were given the full 40 oz. daily:

> "Subjects who had previously stated that they had no specific craving for alcohol drank their beverage as quickly as possible and shortly thereafter felt that more alcohol was necessary to make them feel comfortable. On the third day of the 40 oz. dosage, subjects removed the clock from the wall so that they would not be able to see how long they had to wait before their drink.... Many complained that they could not continue the experiment unless they were given the increased ration of alcohol which they requested."

That remarkably vivid account might in retrospect be read as a description of an experimentally-induced reinstatement of the dependence syndrome. It offers evidence to support the view that reinstatement is more than a myth. From the data that are provided it is not, however, possible to determine the covariance between reinstatement of craving and reappearance of withdrawal symptoms, and even if one were to establish that connection in statistical terms, causality would still be elusive.

Gross *et al.* (1975a) examined slow-

wave sleep as a "carry-over of functional tolerance and physical dependence." Their work represented a commendable attempt to bridge animal and human research but, unfortunately, this project was not taken beyond a preliminary report. Topham (1983) followed 48 hospital-treated alcoholics for six months, at the end of which time 19 had relapsed. The speed of first appearance of withdrawal symptoms after relapse was related to base-line measures of dependence as rated on the SADQ (Stockwell *et al.*, 1979). This too was only preliminary work and needs to be replicated, but it was unusual in that it employed a prospective design.

A further crumb of potentially relevant information comes from data provided by Stockwell *et al.* (1979) in their paper describing the factor structure of the SADQ. The questionnaire was applied to 104 subjects recruited from an alcoholism treatment service. The 'reinstatement' scale correlated at 0.69 with the total SADQ score. What might thus be inferred from those data and from other material in the original report is firstly that reinstatement is a graded phenomenon, and secondly that it is one that is in a mathematical sense integral to dependence.

What implications?

If it is accepted that there is sufficient evidence for the existence of some type of carry-over effect, the necessary next step would be to try to identify the mechanisms responsible for such a phenomenon. LeBlanc & Cappell (1975) have argued that carry-over might firstly and quite simply be related to a further stepwise increment in a level of tolerance and dependence that had as yet not decayed by the time of the second exposure — the basic phenomenon for that kind of incremental effect would be no different from that of the primary phenomenon. The present concern must however be more with LeBlanc & Cappell's second type of situation and the mechanisms that may bear on the organism's capacity to demonstrate a 'memory for addiction' as judged by response to a further challenge exposure to alcohol, administered after all outward evidence of tolerance or withdrawal has decayed. Here LeBlanc & Cappell usefully dissect the problem in the following terms:

> "This effect may take at least three forms: (i) the *rate* at which tolerance and dependence are re-elicited may be increased; (ii) the maximum attainable *level* of tolerance and dependence may be increased; (iii) both... may be increased. This model assumes some relatively durable change in the organism resulting from a period of drug exposure...."

The phase 'some relatively durable change' neatly points to the center of that mystery. Crucial issues to be addressed will include the extent to which this

phenomenon can be related to disturbances in biological systems, or, alternatively, be analysed and understood in learning terms. It may of course be exactly this kind of question that eventually produces a profitable meeting point between biological and psychological ideas.

Meanwhile, it may be useful tentatively to list some of the ways in which the postulated 'addiction memory' may have consequences for the patient and for treatment strategies.

1. If 'addiction memory' exists as a strong and irreversible "change in the organism" affecting a proportion of patients, the previously highly-dependent subject may, as a consequence of this memory, be courting reinstatement of withdrawal symptoms if he ever drinks again, or drinks above some threshold.

2. If, however, the memory mechanisms exist but are for some patients of only limited duration, such patients should be able to drink without elevated danger of reinstating withdrawal if they first establish an appropriate period of abstinence. How to determine which patients would demonstrate irreversible 'addiction memory' or alternatively, would reveal a decay in this memory (over longer or shorter time), would then become a substantially important clinical question. There is an echo here of the quite widely held clinical belief that a 'normal drinking' goal is best pursued only after some prior period of abstinence.

3. If a 'memory' exists which carries significantly adverse implications for the patient's welfare, an important research issue is then the development of methods for suppressing, reversing or eliminating the memory trace. If one believes that the mechanisms involved are in part to be conceived within a learning framwork, then a programme of 'unlearning' based on psychological principles, would be the favored approach (Hodgson & Rankin, 1976). The possibility of developing pharmacological treatments that are targeted in a scientifically informed fashion at dependence mechanisms have as yet only received very preliminary attention (Naranjo & Sellers, 1985).

WITHDRAWAL: WHAT REINFORCING CONSEQUENCES RESULT FOR DRINKING BEHAVIOR?

There is a school of thought that might wish to answer the question raised by this heading with a refreshingly direct phrase such as 'Not a jot.' Adherents to that school regard withdrawal symptoms as epiphenomena. They will point out that the habit of excessive drinking was acquired long before withdrawal distress first appeared. To suggest that these come-lately symptoms perpetuate the drinking is, they will argue, to put the cart before the horse. A powerful and sufficient conjunction of reasons existed for the drinking before the first withdrawal

experience and will remain as sufficient to drive the drinking, shakes or no shakes. The alternative view is that whatever the initial determinants of drinking, once withdrawal symptoms have been established, their relief or avoidance will provide additional cues to reinforce alcohol-seeking behavior.

The research evidence: a series of important reviews

Let's examine what has been said about this issue by a series of authoritative reviews that between them span almost 20 years. Firstly, here is a quotation from an influential early review by Kalant *et al.* (1971):

> "At some point in the drinking history the symptoms arising from the development of physical dependence can become an important contributing etiological factor to the extent that drinking may become a response to the threat of a withdrawal reaction."

Mello (1972) referred particularly to human experiments conducted by Mendelson and herself, and concluded:

> "Once the addictive drug consumption pattern has been re-established, the role of physical dependence in maintaining the pattern is less clear for the alcoholic than for the narcotic user. If avoidance of withdrawal signs were the factor which motivates an alcoholic to continue drinking, then he should consistently taper his drinking when forced to stop by an external constraint. Moreover, it would be expected that during a drinking spree, he should drink consistently enough to avoid partial withdrawal signs and symptoms. Despite the attendant discomfort, the alcoholic does not consistently taper his drinking when forced to stop by an external constraint."

Next in this sequence comes a review by Cicero (1980):

> " ...there seems to be a general acceptance of the premise that animals and humans will drink to forestall the development of a withdrawal syndrome after a period of chronic exposure. This is a reasonable assumption and most investigators would agree that at least some alcoholics do, indeed, state that they drink alcohol to prevent the withdrawal syndrome. However, it is equally clear that many humans will voluntarily undergo a withdrawal syndrome or periods of abstinence.... Thus, withdrawal behavior clearly does not dictate, in any absolute fashion, that alcohol must be consumed;

> the situation is much more complex than this...."

Dealing with animal research, the same author (Cicero, 1980) went on to argue:

> "While there are several reports in the literature which seem to demonstrate that animals will drink more alcohol during withdrawal than under normal conditions... there are many more reports that have failed to establish this relationship."

Cappell & LeBlanc (1981) also reviewed the animal work and concluded:

> " ...the importance of tolerance and physical dependence must be placed in a reasonable perspective. We conclude that there is some evidence that both of these can influence the probability that a drug will be self-administered. At the same time it is clear that vast variations in drug intake can be achieved by manipulations that are totally orthogonal to tolerance and physical dependence."

Referring again to the animal literature, Meisch (1984) favored a more agnostic position:

> "There have been no animal studies in which alcohol was first established as a reinforcer and then the effects of physiological dependence on alcohol intake determined.... Thus it is not known whether alcohol is a better reinforcer when organisms are physiologically dependent."

Stolerman & Goldberg (1986) argued that the focus should today be more on the primary reinforcing effects of drugs rather than withdrawal as reinforcer. Not all abused drugs produce withdrawal symptoms. They conclude that "it is improbable that understanding withdrawal symptoms will be sufficient for a complete understanding of dependence."

Other discussion of these issues has been given by Rigter & Crabbe (1984) and by Wilkins & Gorelick (1986), while the most recent review bearing on this question comes from Grant *et al.* (1990). These authors state:

> "Several studies have shown that the avoidance of withdrawal symptoms after chronic ethanol self-administration was not sufficient for rats... monkeys... or humans... to continue ethanol consumption.... It is possible that physical dependence reflects the extent of alcohol abuse but does not promote abuse... the data suggest that the role of physical dependence in self-administration appears to be minimal."

Addressing the dilemma

The scientific evidence as judged by Grant *et al.* (1990) points toward a negative conclusion and earlier reviews of this literature offer rather diverse judgements but are generally lukewarm toward granting withdrawal symptoms any very salient role as reinforcers. The dilemma lies in the fact that the clinical world would often probably conclude something rather different. Clinicians would be reluctant to dismiss the story of drinking to relieve withdrawal that is contained in the patient transcript quoted earlier in this paper as irrelevant to the understanding of reinforcement. They hear patients talking about the imperative need to reach for a morning drink and the wonderful relief from withdrawal distress that results from the first drink of the day. With back driven against the wall, the clinician who is told that research workers do not believe that withdrawal can become a significant or even at times a salient factor in the field of influences contributing to the learning processes that underly dependence, may even be tempted to the rejoinder "Now you see it now you don't — *again*." Here to sharpen the dilemma is a further quotation from the same patient transcript:

> "The reason I had to have those days off work was... there were days when I was worse than ever. As I was drinking in the morning now, by about 10 a.m. I was really drunk, and it used previously to take me till about 4 p.m. How will I ever do any work?"

That man is in effect saying that withdrawal symptoms led him to drink in the morning, morning drinking shifted the *pattern* of his alcohol intake so that he was drunk earlier in the day, and morning drunkenness imperilled his capacity to hold down a job.

How do we balance that type of story against the assertion that "the role of physical dependence in self-administration appears to be minimal?" It would, of course, be arrogant to dismiss the research evidence, but equally it would be ill-judged to spurn the clinician's insistence that the matter is not yet settled and that there are still some tantalizing mysteries. Here are some of the issues that need to be addressed if we are to move forward:

1. What is the relationship between *intensity* of withdrawal and its potency as reinforcer? Research has certainly established that withdrawal is a graded rather than an all-or-none disturbance and yet much of the literature debates the question of whether withdrawal reinforces alcohol-seeking as if we were dealing with an entity that comes in one size only. It must be noted, for instance, that the other-wise seminal work of Mello & Mendelson (Mendelson & LaDou, 1964a and b; Mendelson *et al.*, 1964; Mendelson & Mello, 1966; Mello & Mendelson, 1970 and 1972) offers little detail on witdrawal severity.

2. When examining the significance of withdrawal as a reinforcer it is in addition unsatisfactory to approach the question only in terms of a global withdrawal severity. It may be necessary to consider the reinforcing potential of separate elements within the total syndrome. Both Gross *et al.* (1973) and Hershon (1977) have shown that affective and physical components of this syndrome can be differentiated. Is it the morning tremor, nausea and sweating, or the morning anxiety and depression that motivated the morning drink, or some integrative function of the two dimensions?

3. Rather than assuming that withdrawal has any real significance as a signal disarticulated from a subject's prior experience of withdrawal and repeated withdrawal relief drinking, should research be directed at the degree to which the previous scheduling of escape or avoidance learning may affect the reinforcing value of withdrawal? Not only conditioning (Wikler, 1961; Ludwig *et al.*, 1974; MacRae *et al.*, 1987), but also cognitive theories (Abrams & Niaura, 1987; Wilson, 1987) need to be harnessed to this problem.

4. There would also today be great advantage in moving beyond an undifferentiated question formed only in terms of whether withdrawal has consequences for holistically conceived 'drinking behavior' and examine instead its relationship to the separate behavioral elements that together contribute to an individual's drinking style. An overall question of the type 'Is withdrawal a reinforcer of drinking behavior' might be useful as a starting point, but it is too crude a formulation to provide much leverage for future research. The most important step toward effecting a reconciliation between science and what patients are telling us may indeed relate to a refocusing of the question toward the influence of withdrawal on the *patterning* of drinking. The question of 'consequence' should be dismantled into a dissection of the potential influence of withdrawal not only on the total alcohol intake over, say, any given 24 hours, but its influence also on the pattern and scheduling of alcohol intake over the given period. What is the significance of withdrawal for the shape of the 24-hour blood alcohol curve? Interesting evidence for a possible relationship between withdrawal intensity and drinking style is to be found in a report by Hershon (1977). Direction of causality cannot be assumed, but Hershon described a significant statistical relationship between scores on both affective and physical disturbance dimensions, and positive answers to items such as 'drinking within an hour of rising', 'drinking to help to get to sleep', 'need to drink regularly during the day' and 'not able to control the amount consumed.'

5. Why do subjects who have been made dependent on alcohol in an experimental situation interrupt their drinking and thus

expose themselves voluntarily to the distress of withdrawal? That this may happen was a crucial finding in the work reported by Mello & Mendelson (1970). This result has on occasion been used as a leading argument against the contention that withdrawal is an important stimulus for drinking. At the extreme there is the fact of 'bout drinking.' What may seem remarkable in relation to this particular question is in fact the absence to date of any systematic research that seeks to elicit the explanations that might be given by subjects or patients themselves as to the reasons for such intermittency. Mello & Mendelson (1970) did not, for instance, report any formal debriefing of their patients on this topic and it is difficult to find any clinical research that has asked patients such a question as 'Why do you interrupt or stop your drinking?' It could be conjectured that for a patient who is operating within a pay-off matrix that determines whether or not he continues to drink (Orford, 1970), the moment may arrive when the perceived benefits of drinking (withdrawal relief) will be outweighed by the perceived benefits of not drinking (not being sick tomorrow, avoiding gastritis, and so on). To approach the analysis in those terms one must, of course, first escape from the constraint of believing that what has to be tested is whether or not withdrawal drives drinking in some inexorable and absolute fashion. Behavior is always likely to be determined by a field of influence rather than by any one vector. There is some evidence to suggest that in an experimental situation monkeys will show an interrupted pattern of drinking (Karocy *et al.*, 1978).

LISTENING AND THE LABORATORY

If we conclude that Trotter's observation that the hands of some inebriates tremble in the morning but become steady as the dose is repeated can still point toward an important natural mystery, what are the strategies that will now offer further research progress? As this review has shown, there is much on which to build.

Here are two brief suggestions. Firstly, there is a need to strengthen the input to the total effort that comes from a type of research that has no precise designation but which for shorthand one might call 'listening and looking.' The research community needs to set about the task of strengthening an investigative tradition that is rooted in listening to what patients have to tell us and observing what they have to show. That tradition has always made a contribution, but its methodologies stand today in need of enhancement. It is wrong to dismiss this type of approach as 'soft.' If researchers are to do a better job of making sense out of difficult raw material, there is a need to develop standard interview methods and questionnaires that are today in many areas not available. There is also the possibility of exploiting methods for qualitative textual analyses that can be applied to

patients' accounts of what happened and why it happened. Furthermore, to refer to the informants as 'patients' is inadequate, and there is a need to work more often with non-clinical samples. Finally, if it is accepted that many of the key questions in this area relate to what happens over time, the capacity to stay close to our subjects and conduct longitudinal research needs to be developed.

The second suggestion follows on directly from the first. The future might be seen as demanding a closer working liaison than has so far been established between a new wave of listening and observational research on the one hand, and experimental and laboratory research on the other. Research must reach the point where the power of different methods and theories can be brought sharply to bear on shared questions — that has often been the talk, but seldom the actuality.

At the end of this discussion it seems only proper to return to Keats. "Axioms are not axioms," he said, "until they are proved upon our pulses." Science is not science, he seems to be saying, until it is tested against what our tremulous patients have to tell us, if only we will listen.

ACKNOWLEDGEMENTS

I am grateful to the Addiction Research Foundation for the invitation that led to the preparation of this paper. Mrs. Patricia Davis provided expert secretarial support.

This lecture was also published in the *British Journal of Addiction* 85(4):447-461, 1990.

REFERENCES

Abrams, D. B. & Niaura, R. S. (1987) Social learning theory, in: H. T. Blane & K. E. Leonard (Eds.) *Psychological Theories of Drinking and Alcoholism*, Chapter 5, pp. 131-178 (New York, Guildford Press).

Baker, T. B. & Cannon, D. S. (1979) Potentiation of ethanol withdrawal by prior dependence, *Psychopharmacology*, 60, pp. 105-200.

Bowman, K. A. & Jellinek, E. M. (1942) Alcoholic mental disorders, in E. M. Jellinek (Ed.) *Alcohol Addiction and Chronic Alcoholism*, 2, pp. 8-172 (New Haven, Yale University Press).

Branchey, M., Rauscher, G. & Kissen, B. (1971) Modification in the response to alcohol following the establishment of physical dependence, *Psychopharmacology*, 22, pp. 314-322.

Cappell, H. & LeBlanc, A. E. (1979) Tolerance to and physical dependence on ethanol. Why do we study them? *Drug and Alcohol Dependence*, 4, pp. 15-31.

Cappell, H. & LeBlanc, A. E. (1981) Tolerance and physical dependence. Do they play a role in alcohol and drug self-administration? in: Y. Israel, F. B. Glaser, H. Kalant, R. E. Popham, W. Schmidt & R. G. Smart (Eds.) *Research Advances in Alcohol and Drug Problems*, Vol. 6, pp. 159-161 (New York, Plenum Press).

Cicero, T. J. (1980) Alcohol self-administration, tolerance and withdrawal in humans and animals. Theoretical and methodological issues, in: H. Rigter & J. C. C. Rabbe (Eds.) *Alcohol Tolerance and Dependence*, 1, pp. 1-51 (New York, Elsevier).

Deutsch, J. A. & Walton, N. Y. (1977) A rat alcoholism model in a free choice situation, *Behavioural Biology*, 19, pp. 349-360.

Edwards, G. & Gross, M. M. (1976) Alcohol dependence: a provisional description of a clinical syndrome, *British Medical Journal*, pp. 1058-1061.

Ellis, F. W. & Pick, J. R. (1971) Ethanol intoxication and dependence in rhesus monkeys, in: N. K. Mello & J. R. Mendelson (Eds.) *Recent Advances in Studies of Alcoholism: an interdisciplinary symposium*, pp. 401-412 (Washington, DC, US Government Printing Office).

Falk, J. L. (1983) Drug dependence: myth or motive? *Pharmacology, Biochemistry and Behaviour*, 19, pp. 385-391.

Fingarette, H. (1988) *Heavy Drinking: the Myth of Alcoholism as a Disease* (Berkeley, University of California Press).

Friedman, H. J. (1980) Assessment of physical dependence on and withdrawal from ethanol in animals, in: H. Rigter & J. C. Crabbe (Eds.) *Alcohol Tolerance and Dependence*, Chapter 4, pp. 99-355 (New York, Elsevier).

Goldstein, D. B. (1972) Relationship of alcohol dose to intensity of withdrawal signs in mice, *Journal of Pharmacology and Experimental Therapeutics*, 180, pp. 203-215.

Goldstein, D. B. (1974) Rates of onset and decay of alcohol physical dependence in mice, *Journal of Pharmacology and Experimental Therapeutics*, 190, pp. 377-383.

Goldstein, D. B. & Arnold, W. (1976) Drinking patterns as predictors of alcohol withdrawal reactions in DBA/23 mice, *Journal of Pharmacology and Experimental Therapeutics*, 199, pp. 408-414.

Grant, K. A., Hoffman, P. & Tabakoff, B. (1990) Neurobiological and behavioural approaches to tolerance and dependence, in: G. Edwards & M. Lader (Eds.) *The Nature of Dependence,* pp. 135-162 (Oxford University Press).

Gross, M. M., Lewis, E. & Nagarajan, M. (1973) An improved quantitative system for assessing the acute alcoholic psychoses and related states (TSA and SSA), in: M. M. Gross (Ed.) *Alcohol Intoxication and Withdrawal: Experimental Studies. Advances in Experimental Medicine and Biology*, Vol. 30, pp. 365-376 (New York, Plenum Press).

Gross, M. M., Hastey, J. M., Lewis, L. & Young, N. (1975a) Slow wave sleep and carry-over of functional tolerance and physical dependence in alcoholics, in: M. M. Gross (Ed.) *Alcohol Intoxication and Withdrawal II. Advances in Experimental Medicine and Biology,* Vol. 59, pp. 477-493 (New York, Plenum Press).

Gross, M. M., Lewis, E., Best, S., Young, N. & Feuer, L. (1975b) Quantitative changes of signs and symptoms associated with acute alcohol withdrawal: incidence, severity and circadian effects in experimental studies of alcoholics, in: M. M. Gross (Ed.) *Alcohol Intoxication and Withdrawal II. Advances in Experimental Medicine and Biology*, Vol. 59, pp. 615-643 (New York, Plenum Press).

Haggard, H. W. (1945) The physiological effects of large and small amount of alcohol, in: *Alcohol, Science and Society. Twenty-nine Lectures with Discussions as given at the Yale Summer School of Alcohol Studies*, Lecture 5, pp. 59-72 (New Haven, *Quarterly Journal of Studies on Alcohol*).

Hare, F. (1910) The withdrawal of narcotics from habituees, *British Journal of Inebriety*, 8, pp. 86-90.

Hershon, H. I. (1977) Alcohol withdrawal symptoms and drinking behaviour, *Journal of Studies on Alcohol*, 38, pp. 953-971.

Hodgson, R. J. & Rankin, H. (1976) Modification of excessive drinking by cue exposure, *Behaviour Research and Therapy*, 14, pp. 305-307.

Huss, M. (1849) *Alcoholismus chronicus eller chronisk* (Stockholm, Alkoholssjukdom).

Isbell, H., Fraser, H. F., Wikler, A., Belleville, R. E. & Eiserman, A. J. (1955) An experimental study of the etiology of 'rum fits' and delirium tremens, *Quarterly Journal of Studies on Alcohol*, 16, pp. 1-33.

Jellinek, E. M. (Ed.) (1942) *Alcohol Addiction and Chronic Alcoholism* (New Haven, Yale University Press).

Jellinek, E. M. (1952) Phases in the drinking history of alcohlics, *Quarterly Journal of Studies on Alcohol*, 7, pp. 1-88.

Jellinek, E. M. (1960) *The Disease Concept of Alcoholism* (New Brunswick, Hillhouse).

Joliffe, N. (1945) Alcohol and Nutrition. The Diseases of Chronic Alcoholism, in: *Alcohol Science and Society, Twenty-nine Lectures with Discussions as given at the Yale Summer School of Alcohol Studies*, Lecture 6, pp. 73-92 (New Haven, *Quarterly Journal of Studies on Alcohol*).

Kalant, H. (1973) Biological models of alcohol tolerance and physical dependence, in: M. M. Gross (Ed.) *Alcohol Intoxication and Withdrawal: Experimental Studies. Advances in Experimental Medicine and Biology*, Vol. 35, pp. 3-13 (New York, Plenum Press).

Kalant, H., LeBlanc, A. E. & Gibbins, R. J. (1971) Tolerance to, and dependence on, ethanol, in: Y. Israel & J. Mardones (Eds.) *Biological Basis of Alcoholism*, Chapter IX, pp. 235-269 (New York, Wiley-Interscience).

Kalant, H., LeBlanc, A. E., Gibbins, R. J. & Wilson, A. (1978) Accelerated development of tolerance during repeated cycles of ethanol exposure, *Psychopharmacologia*, 60, pp. 59-65.

Karoly, A. J., Winger, G., Ikomi, F. & Woods, J. H. (1978) The reinforcing property of ethanol in the rhesus monkey II. Some variables related to the maintenance of intranervous ethanol-reinforcing responding, *Psychopharmacologia*, 58, pp. 19-25.

Keats, J. (1817) Letters, 32.

Kendell, R. E., De Roumanie, M. and Ritson, E. B. (1983) Effect of economic changes on Scottish drinking habits 1978-82, *British Journal of Addiction*, 78, pp. 365-379.

LeBlanc, A. E. & Cappell, H. D. (1975) Historical antecedents as determinants of tolerance and dependence upon psychoactive drugs, in: H. D. Cappell & A. E. LeBlanc (Eds.) *Biological and Behavioural Approaches to Drug Dependence*, pp. 43-51 (Toronto, Addiction Research Foundation).

Ludwig, A. M., Wikler, A. & Stark, L. H. (1974) The first drink; psychobiological aspects of craving, *Archives of General Psychiatry*, 30, pp. 539-547.

Marlatt, G. A. (1985) Relapse prevention: theoretical rationale and overview of the model, in: G. A. Marlatt & J. R. Gordon (Eds.) *Relapse Prevention*, Chapter 1, pp. 3-70 (New York, Guildford Press).

Meisch, R. A. (1984) Alcohol self-administration by experimental animals, in: R. G. Smart, H. D. Cappell, F. B. Glaser, Y. Israel, H. Kalant, R. E. Popham, W. Schmidt & E. M. Sellers (Eds.)

Research Advances in Alcohol and Drug Problems, Vol. 8, Chapter 2, pp. 23-45 (New York, Plenum Press).

MacRae, J. R., Scoles, M. T. & Siegel, S. (1987) The contribution of Pavlovian conditioning to drug tolerance and dependence, *British Journal of Addiction*, 83, pp. 371-380.

Medical Research Council (1938) *Alcohol, its Action on the Human Organism*, 3rd edn. revised [1st edn. 1918] (London, HMSO).

Mello, N. K. (1972) Behavioral studies of alcoholism, in: B. Kissin & H. Begleiter (Eds.) *Biology of Alcoholism*, Vol. 2, Chapter 9, pp. 219-291 (New York, Plenum).

Mello, N. K. & Mendelson, J. H. (1970) Experimentally induced intoxication in alcoholics: a comparison between programmed and spontaneous drinking, *Journal of Pharmacology and Experimental Therapeutics*, 173, pp. 101-116.

Mello, N. K. & Mendelson, J. H. (1972) Drinking patterns during work-contingent and non-contingent alcohol acquisition, *Psychosomatic Medicine*, 34, pp. 139-164.

Mendelson, J. H. (Ed.) (1964) Experimentally induced chronic intoxication and withdrawal in alcoholics, *Quarterly Journal of Studies on Alcohol*, Supp. 2.

Mendelson, J. H. & LaDou, J. (1964a) Background and experimental design, in: J. H. Mendelson (Ed.) Experimentally induced chronic intoxication and withdrawal in alcoholics,lics, *Quarterly Journal of Studies on Alcohol*, Suppl. 2, Part 1, pp. 1-13.

Mendelson, J. H. & LaDou, J. (1964b) Psychophysiological findings, in: J. H. Mendelson (Ed.) Experimentally induced chronic intoxication and withdrawal in alcoholics, *Quarterly Journal of Studies on Alcohol*, Supp. 2, Part 2, pp. 14-39.

Mendelson, J. H., LaDou, J. & Solomon, P. (1964) Psychiatric findings, in: J. H. Mendelson (Ed.) Experimentally induced chronic intoxication and withdrawal in alcoholics, *Quarterly Journal of Studies on Alcohol*, Suppl. 2, Part 3, pp. 40-52.

Mendelson, J. H. & Mello, N. K. (1966) Experimental analysis of drinking behavior of chronic alcoholics, *Annals of the New York Academy of Science*, 33, pp. 828-845.

Naranjo, C. A. & Sellers, E. M. (1985) *Psychiatric Advances in New Psychopharmacological Treatments for Alcoholism* (Amsterdam, Excerpta Medica).

Nathan, P. E., Titler, N. A., Lowenstein, L. M., Solomon, P. & Rossi, A. M. (1970) Behavioral analysis of chronic alcoholism, *Archives of General Psychiatry*, 22, pp. 419-430.

Orford, J. (1970) Aspects of the relationship between alcohol and drug abuse, in: L. G. Kiloh & D. S. Bell (Eds.) *29th International Congress on Alcoholism and Drug Dependence*, pp. 130-144 (Chatswood, Butterworths (Australia)).

Orford, J. (1985) *Excessive Appetites: A Psychological View of Addictions* (Chichester, John Wiley).

Poldrugo, F. & Snead, O. C. (1984) Electroencephalographic and behavioural correlates in rats during repeated ethanol withdrawal syndrome, *Psychopharmacology*, 83, pp. 140-146.

Rigter, H. & Crabbe, J . C. (Eds.) (1980) *Alcohol Tolerance and Dependence* (New York, Elsevier).

Stockwell, T., Hodgson, R., Edwards, G., Taylor, C. & Rankin, H. (1979) The development of a questionnaire to measure severity of alcohol dependence, *British Journal of Addiction*, 74, pp. 79-87.

Stolerman, I. P. & Goldberg, S. R. (1986) Introduction: Brief history and scope of behavioral approaches to dependence, in: S. R. Goldberg & I. P. Stolerman (Eds.) *Behavioral Analysis of Drug Dependence*, Chapter 1, pp. 1-7 (Orlando, Academic Press).

Topham, A. (1983) Alcohol dependence and craving, *Ph.D. Thesis* (University of London).

Trotter (1804) An Essay, *Medical, Philosophical and Chemical on Drunkenness and its Effects on the Human Body* (London: Longman & Rees).

Turner, T. B. (1980) Clinical aspects of ethanol tolerance and dependence, in: H. Rigter & J. C. Crabbe (Eds.) *Alcohol Tolerance and Dependence*, Chapter 16, pp. 395-424 (New York, Elsevier).

Victor, M. & Adams, R. D. (1953) The effect of alcohol on the nervous system, *Research Publications of the Association for Research on Nervous and Mental Disease*, 32, pp. 526-573.

Walker, D. W. & Zornetzer, S. F. (1974) Alcohol withdrawal in mice: electroencephalographic and behavioral correlates, *Electroencephalography and Clinical Neurophysiology*, 36, pp. 233-244.

Wallgren, H. & Barry, H. (1970) *Actions of Alcohol Volume II. Chronic and Clinical Aspects* (Amsterdam, Elsevier).

Wikler, A. (1961) On the nature of addiction and habituation, *British Journal of Addiction*, 57, p. 16.

Wilkins, J. N. & Gorelick, D. A. (1986) Clinical neuroendocrinology and neuropharmacology of alcohol withdrawal, in: M. Galanter (Ed.) *Recent Developments in Alcoholism*, Vol. 4, Chapter 11, pp. 241-263 (New York, Plenum).

Wilson, G. T. (1987) Cognitive processes in addiction, *British Journal of Addiction*, 82, pp. 343-354.

World Health Organization (1951) *Expert Committee on Mental Health. Report of the First Session of the Alcoholism Subcommittee*, WHO Technical Report Series No. 42 (Geneva, WHO).

World Health Organization (1952) *Expert Committee on Drugs Liable to Produce Addiction. Third Report*, WHO Technical Report Series No. 57 (Geneva, WHO).

World Health Organization (1954) *Expert Committee on Alcohol, First Report*, WHO Technical Report Series No. 94 (Geneva, WHO).

World Health Organization (1955) *Alcohol and Alcoholism. Report of an Expert Committee*, WHO Technical Report Series No. 94 (Geneva, WHO).

Yale University Center of Alcohol Studies (1945) *Alcohol, Science and Society* (New Haven: *Quarterly Journal of Studies on Alcohol*). [Reprinted 1971 by Greenwood Press, Westport.]

◆

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

Evolving Knowledge on Genetic Factors in Alcoholism

C. ROBERT CLONINGER

ERIC J. DEVOR

DEPARTMENT OF PSYCHIATRY

AND GENETICS

WASHINGTON UNIVERSITY

SCHOOL OF MEDICINE

ST. LOUIS, MISSOURI

C. ROBERT CLONINGER

C. Robert Cloninger is the Wallace Renard Professor at Washington University School of Medicine in St. Louis. He is also head of the department of Psychiatry and director of the Mental Health Clinical Research Centre in Psychiatric Genetics and Family Studies. Dr. Cloninger is president-elect of the American Psychopathological Association, on the board of directors of the Research Society of Alcoholism, a member of the Institute of Medicine of the National Academy of Sciences, and associate editor of the journals *Human Heredity* and *Genetic Epidemiology*. His current work includes studies of the genetic epidemiology of psychiatric disorders; refining the assessment and subtyping of alcoholism, personality disorders and schizophrenia; longitudinal studies of adoptees at risk for alcoholism and related disorders; and molecular genetic linkage studies of the inheritance of alcoholism.

The study of the genetics of alcoholism is at a watershed. Since ancient times, the familial nature of alcoholism has been recognized. This familial aggregation has been consistently documented in studies throughout the last century. However, alcoholism seems not to be inherited in a simple Mendelian manner, and such familial aggregation may be due to genetic influences, familial environmental influences, or both. Initially the familial aggregation was attributed to social and environmental factors, and the importance of genetic factors was discounted or minimized with little critical testing. Then between 1960 and 1980 studies of twins and adoptees provided increasing evidence for the importance of genetic factors. However, the results concerning the relative importance of genetic and environmental factors were inconsistent. Then in 1981 evidence for genetic heterogeneity in a large adoption study was presented. The inheritance of one type of alcoholism involves the interaction of both genetic and home environmental factors; whereas, the second type is highly heritable in men regardless of home environmental factors. This heterogeneity explained the earlier inconsistencies about the inheritance of alcohol-related behavior disorders.

The clinical and genetic heterogeneity observed in adoptees was later confirmed in other family studies and extended by the demonstration of associated differences in heritable neurochemical and neurophysiological factors. Most recently, a comprehensive model of neurogenetic adaptive mechanisms underlying susceptibility to alcoholism has been presented. This has stimulated efforts, now underway, to identify specific genetic susceptibility factors in alcoholism using linkage mapping. Thus the evolution of our knowledge about genetic factors in alcoholism has been influenced by contemporary theories about human behavior and by the availability of research tools that allow tests of alternative hypotheses. In this paper, we will first review the evolving state of knowledge and expert opinion. Then we will discuss contemporary research and potential future directions for clarifying the role of genetic factors in alcoholism.

VIEWS BEFORE THE TWENTIETH CENTURY

The familial aggregation of alcoholism has been recognized anecdotally since ancient times. Both Aristotle and Plutarch believed that alcoholics "begat" other alcoholics (Goodwin, 1987). In the Middle Ages, Chaucer repeatedly used the word "dronkelew" and suggested that dependence on alcohol was a familial mental illness (Goodwin & Guze, 1989).

However, systematic family studies of alcohol abusers were not reported until the late nineteenth century. In 1879, Long attributed 21 per cent of all inherited disease in the children of alcoholics to alcoholism in their parents. In 1909,

Crothers found that 70 per cent of more than 4,000 alcoholics he had treated in the prior 35 years had reported relatives who drank heavily. In the past 80 years more than 100 studies have confirmed Crothers' finding that alcoholism is familial (Goodwin, 1976).

FAMILY STUDIES IN PAST FOUR DECADES

The familial aggregation of alcoholism is one of the most consistent findings in medicine. Cotton (1979) summarized data on the families of 6,251 alcoholics and 4,083 non-alcoholics gleaned from 39 studies conducted over the past four decades. On average, almost one-third of any sample of alcoholics had at least one parent with an alcohol abuse problem. In different studies, however, the proportion of alcoholics with no alcoholic parent varied from 45 to 80 per cent, but the general finding of familial aggregation was consistently supported.

Before 1960 the familial aggregation of alcoholism was widely attributed to social and environmental factors in the home despite the consistency of inheritance regardless of historical period and culture. Roe and Burks (1945) found that good foster placement was associated with no alcohol abuse in all but one of 27 children of alcoholic biological parentage and in all but one of 22 children of normal biological parentage. They found that none of 21 adopted-away sons of alcoholics had drinking problems as adults. This reinforced the popular assumption that psychopathology was caused by early home environmental problems and could be prevented by good home placement. This study was later criticized because of its vague and weak definition of alcoholism and small sample size. For example, none of the parents studied by Roe and Burks had been treated for alcoholism. Also, the alcohol abusers' children whom they studied were more often placed in rural areas where drinking was infrequent than were the children of nonalcoholics. As late as 1960, Jellinek recognized that alcoholism was familial but claimed that there was no evidence for a genetic factor. This conclusion was justified at the time. However, the next two decades would confirm the logical fact that the absence of proof did not constitute a proof of the absence of genetic influences on susceptibility to alcoholism.

TWIN AND ADOPTION STUDIES FROM 1960-1980

Kaij launched the modern era of systematic twin and adoption studies of alcoholism in 1960. Several studies of recurrent alcohol abuse in Swedish men are summarized in **Table 1**, including the findings of Amark (1951) in ordinary sibs, Kaij (1960) in twins, Bohman (1978) in adopted children, and Kaij and Dock

Table 1. Family Studies of Recurrent Alcohol Abuse in Swedish Men

SOURCE		RELATIONSHIP	PROBANDWISE CONCORDANCE N	%
Kaij	(1960)	MZ twins	27	70
Kaij	(1960)	DZ twins	60	33
Amark	(1951)	Singleton sib	349	31
		• No alcoholic parent	252	17
		• One alcoholic parent	97	33
Bohman	(1978)	Adopted-away sons	50	20
Kaij and Dock	(1975)	Grandsons	270	12
Census	(1968)	General population		7

(1975) in grandsons. Kaij (1960) found that 70 per cent of the co-twins of an alcoholic monozygotic twin were themselves alcoholic, compared to only 33 per cent of the co-twins of an alcoholic dizygotic twin. A comparable risk for recurrent alcohol abuse in men in the general Swedish population was only seven per cent. Since monozygotic twins are genetically identical and dizygotic twins share only half of their genes, this suggested that genetic factors were important. Furthermore the observations in other classes of relatives showed that the risk of recurrent alcohol problems was highly correlated with the degree of genetic relatedness.

Kaij's work called into question the assumption that genetic factors could be ignored, but reservations remained about the particular aspects of susceptibility to alcoholism that were heritable. Recurrent alcohol abusers are clinically heterogeneous, so large scale studies were needed to determine what aspects of alcohol abuse were heritable. Accordingly, Partanen and his associates (1966) carried out a large scale study of drinking behavior in Finland.

The heritabilities of four drinking-related quantitative variables were computed in the Finnish twin study. Factor 1 (density or frequency of drinking) and factor 2 (amount of drinking) had heritabilities of 39 per cent and 36 per cent respectively. In contrast, lack of control and social complications showed little evidence of heritability except in younger subjects. However, the first latent root obtained in a canonical correlation analysis showed a heritability of 66 per cent and was determined by high loadings on both amount of drinking and lack of control.

Partanen and associates (1966)

interpreted their findings to show that frequency and amount of drinking show considerable hereditary variation, whereas loss of control and social complications are almost entirely the result of environmental variation. Their interpretation is entirely consistent with the heritabilities reported for the individual factor scores. However, the findings of the canonical correlation analysis suggested the alternative hypothesis that loss of control is also determined by genetic factors but that exposure to continued heavy drinking is required before that genetic influence can be manifest (Cloninger, Reich, Wetzel, 1979; Cloninger & Reich, 1983). If the two sets of hereditary factors operate independently, the appearance of environmental determination of loss of control results in a general sample that has not been selected for heavy drinking. The latter reinterpretation has been clearly confirmed by recent adoption studies (Cloninger *et al.*, 1981), but appears to have been overlooked until 1979.

Following the twin studies of drinking behavior in the 1960s, the next major advance in knowledge occurred on the basis of studies of Danish adoptees reported by Goodwin and associates (1973, 1977a,b). They studied children who had been adopted at an early age away from biological parents who had been hospitalized for alcoholism. They found that chronic alcoholism was four times as common in 55 adopted-away sons of alcoholics as in 78 such sons of nonalcoholics. However, alcoholism in adopted-away daughters of Danish alcoholics was not significantly increased above the incidence in adopted controls. Furthermore, lesser forms of problem drinking were not increased in adoptees of either sex and heavy drinking was actually significantly decreased in the non-adopted sons (Cloninger & Reich, 1983). This suggests that the control group in the Danish studies may include many children of nonalcoholic heavy drinkers. Goodwin and his associates concluded that genetic factors, but not home environment, were important in the development of severe alcoholism.

The Danish adoption studies were admired for their explicit definition of alcoholism and the blind assessments of adoptees without knowledge of the biological parent's drinking status. However, they were criticized for their moderate sample size and limited information about the home environment of the adoptees. The Danish studies were confirmed in some smaller scale American studies of adoptees (Cadoret & Gath, 1978). However, lack of consistency between the Roe and Burks (1945) and the later Danish and American studies, combined with the low heritability of severe alcoholism in the Finnish and some other twin studies, led to continued controversy about the relative importance of genetic and environmental factors. Skeptics could remain doubtful about the magnitude of genetic effects, and many social scientists did remain skeptical. Nevertheless, it seemed likely after the

Danish adoption studies that genetic factors should not be ignored.

ADOPTION STUDIES AFTER 1980: THE SWEDISH STUDIES

In view of the uncertainties about the importance of genetic and social factors in alcoholism, Michael Bohman and his co-workers undertook a large scale adoption study of alcoholism in Sweden. Bohman actually began the study around 1970 with the hypothesis that adoptive placement in relatively stable adoptive homes would eliminate the social heritage that was popularly assumed to cause alcohol abuse. He was aware of both Kaij's and Partanen's results, and was more impressed by the importance of social, rather than genetic, factors.

Bohman's subjects included all the children born to single women in Stockholm from 1913 to 1965. Most analyses have focused on the 862 men and 913 women of known paternity who were born from 1930 to 1949 and adopted by non-relatives at an early age. These individuals were reared during a time when both men and women had ready access to alcohol and systematic registry information about alcohol abuse from the Swedish Temperance Boards. Most of the subjects were separated from their biological relatives in the first months of life (average of four months) and all had their final placement before the age of three years (average of eight months). Information about alcohol abuse, psychopathology, and medical treatment was available for the entire lifetimes of the adoptees and their parents from hospitals, clinics, and several registers that are systematically maintained in Sweden. Identification of alcohol abuse from these sources can identify about 70 per cent of alcoholics; those so identified are representative of alcoholics in general, with no appreciable bias for either type 1 or type 2 alcoholics (Ojesjo, 1980; Kaij, 1970). Such detailed information permitted a population-based study large enough that the adoptees could be divided into subgroups in order to explore the possible genetic heterogeneity of alcoholism.

Earlier studies had suggested that the heritability of alcoholism may be influenced by the differences in the severity of alcoholism. Accordingly, Cloninger, Bohman, and Sigvardsson (1981) divided subjects into mild, moderate, and severe alcohol abusers and non-abusers. Severe alcoholics were individuals with a psychiatric diagnosis of alcoholism, hospital treatment for alcoholism, or more than three registrations for intemperant behavior. Kaij (1960) had found that such criteria reliably identified individuals considered definite alcoholics on detailed clinical interviews. Moderate abusers were individuals with two or three registrations for alcohol abuse but no diagnosis or medical treatment for alcoholism. Kaij had found that such

Table 2. Inheritance of Susceptibility to Alcohol Abuse in the Stockholm Study

ALCOHOL ABUSE IN BIOLOGICAL PARENTS		ALCOHOL ABUSE IN ADOPTEES			
		SONS		DAUGHTERS	
FATHER	MOTHER	N	%	N	%
No	No	571	14.7	577	2.8
Yes	No	259	22.4	285	3.5
No	Yes	23	26.0	29	10.3
Yes	Yes	9	33.3	22	9.1

individuals were problem drinkers but not strictly alcohol dependent. Mild abusers were those subjects with single registrations for alcohol abuse but no diagnosis or medical treatment for alcoholism.

Contrary to Bohman's expectation, alcohol abuse in the adoptive parents was not associated with an increased risk of alcohol abuse in the adoptees, suggesting that imitation of parental drinking was not a major determinant of their alcohol abuse (Cloninger *et al.*, 1981; Bohman *et al.*, 1981). In contrast, alcohol abuse in the biological parents was associated with a substantially increased risk of alcohol abuse in their adopted-away children. The risk of alcohol abuse was studied in the children when they were 23 to 43 years of age. As shown in **Table 2**, the adopted-away sons were more likely to be alcohol abusers if either their biological father or mother had been registered for alcohol abuse than if neither parent had been registered for alcohol abuse. However, the adopted-away daughters were more likely to be alcohol abusers only if their biological mother had been registered for alcohol abuse, not if neither parent or only the biological father had been registered for alcohol abuse (as in the Danish adoption studies). This sex difference suggested that some types of alcohol abuse might be heritable primarily in men, whereas other forms might be heritable in both men and women. This heterogeneity accounted for the failure of the Danish studies to detect genetic influences on alcoholism in women because nearly all the alcoholic parents in the Danish studies were fathers (Bohman *et al.*, 1981). In order to evaluate the possible genetic heterogeneity of alcohol abusers further, the inheritance patterns of mild, moderate, and severe abusers were compared using discriminant analysis. The biological fathers of moderate alcohol abusers were found to differ significantly from those of the parents of the other abusers: their fathers more often had teenage onset of both criminal behavior and severe alcohol abuse. In contrast the biological fathers of

Table 3. Cross-fostering Analysis of Mild and Severe Type 1 Alcoholism

IS GENETIC BACKGROUND TYPE 1?	IS ENVIRONMENTAL BACKGROUND MILD OR SEVERE?	MALE ADOPTEES OBSERVED		
		TOTAL N	% WITH MILD ABUSE	% WITH SEVERE ABUSE
No	–	448	6.5	4.2
Yes	No	237	7.2	6.3
Yes	Mild	91	15.4 *	7.7
Yes	Severe	86	4.7	11.6 *

* Abuse is increased only given both genetic and postnatal predisposition ($p<.05$).

the mild and severe alcohol abusers usually had later adult onset of mild alcohol abuse without associated criminal behavior; this was true whether the adoptee's abuse was mild or severe. Furthermore, the biological mothers of the mild or severe abusers had a substantially increased risk of alcohol abuse; the biological mothers of the type 2 alcohol abusers had no excess compared to the mothers of non-alcoholic adoptees. The only observed differences in the backgrounds of the mild and severe alcohol abusers was in the occupational status of their biological and adoptive parents; severe abuse was more likely to occur in those with predisposition to low occupational status, which was in turn associated with heavier recreational drinking. These findings suggested that there were two patterns of inheritance: Type 1, as seen in both mild and severe alcohol abusers, and type 2, as seen in moderate abusers. In other words, the differences between mild and severe alcohol abusers were largely determined by environmental exposure variables, not genetic factors.

The importance of postnatal environmental factors on severity of type 1 alcohol abuse was examined in a cross-fostering analysis of the 862 adopted men. In a cross-fostering analysis, the risk of disorder in adoptees is considered as a consequence of different combinations of biological parent background and postnatal environmental experience. This is illustrated by **Table 3** for alcohol abuse in the 862 adopted men. The risk of severe abuse doubled in sons who had both type 2 biological parents (as described above) and low occupational status in their adoptive home. The risk of mild abuse likely depended on both the genetic and the environmental background. Likewise, the risk of mild abuse depended on both genetics and the environmental background. Accordingly, type 1 alcoholism has also been called "milieu-limited" alcoholism. This is the type of alcohol abuse that was characteristic of the

Table 4. Cross-fostering Analysis of Type 2 Alcohol Abuse in Men

IS GENETIC BACKGROUND TYPE 2?	IS ENVIRONMENTAL BACKGROUND TYPE 2?	MALE ADOPTEES OBSERVED	
		TOTAL N	% WITH TYPE 2 ABUSE
No	No	567	1.9
No	Yes	196	4.1
Yes	No	71	16.9 *
Yes	Yes	28	17.9 *

* Risk is significantly increased in those with type 2 genetic background compared to others ($p<0.01$).

subjects in Roe and Burk's adoption study.

In contrast, type 2 alcoholism was highly heritable from father to son regardless of external circumstances (**Table 4**). The risk in sons of type 2 fathers was increased about nine-fold compared to the risk in other fathers, including those with either type 1 or no alcoholism. Most subjects in the Danish studies appear to correspond to the type 2 alcoholics in the Swedish studies.

To evaluate this apparent genetic heterogeneity further, predictions were tested in the adopted-away daughters of type 1 and type 2 alcoholics. The background of the daughter's biological parent was classified as it was in the study of the men. The daughters of type 1 alcoholics were predicted to be at increased risk for alcohol abuse because the mothers in these families were often alcohol abusers; the daughters at high risk for type 1 alcoholism were three times as likely to abuse alcohol as those at low risk. The daughters of type 1 alcoholics were not at high risk for other psychopathology. In contrast, the adopted-away daughters of type 2 alcoholic fathers were found to be at higher risk only for somatic anxiety (that is, somatization or frequent disabling physical complaints). Consequently, type 2 alcoholism has been called "male-limited."

RECENT FAMILY AND HIGH-RISK STUDIES

In order to evaluate the generalizability of the Swedish results, Cloninger and Reich and their associates have carried out pedigree analyses of the families of 286 hospitalized alcoholics. Analyses of these pedigrees confirmed that the families of alcoholics were heterogeneous. Two types of families were identified that corresponded closely to those described in the Swedish adoption studies (Gilligan *et al.*, 1987). Furthermore, the men and women in these two types of families had

different patterns of alcohol-related symptoms and personality traits, as summarized in **Table 5** (Gilligan *et al.*, 1988; Cloninger, 1987). The type 1 pattern of inheritance is associated with adult onset of loss of control, binge drinking, guilt feelings about drinking, and alcohol-related liver disorders. Such subjects had anxious or passive-dependent personalities characterized by high harm-avoidance (that is, apprehensive, inhibited, pessimistic and shy), high reward-dependence (that is, sentimental, socially sensitive), and low novelty-seeking (that is, reflective, rigid, stoical and frugal). Such individuals often enjoy the anxiety-reducing effects of alcohol.

In contrast, the type 2 pattern of inheritance was associated with inability to abstain entirely from the euphoriant effects of alcohol and early social problems such as fighting and arrests when drinking. Such subjects often had antisocial personalities characterized by high novelty-seeking (that is, impulsive, exploratory, disorderly and extravagant), low harm-avoidance (that is, confident, uninhibited, optimistic, risk-taking), and low reward-dependent (that is, socially detached, emotionally cool, practical, and independent). Such individuals often actively seek out alcohol and other drugs for their stimulant effects.

In order to test whether these personality traits were antecedents or consequences of alcohol abuse, we analyzed prospective data about adoptees who had been followed from birth to 28 years of age in Sweden (Cloninger *et al.*, 1988). Personality traits of harm avoidance, novelty seeking, and reward dependence were rated when the children were 10 or 11 years of age without knowledge of the adult behavior or that of their parents. These ratings were based on interviews with the school teachers of the children when they were in the fourth grade. Then the children's record of alcohol abuse was evaluated when they were 27 years of age. Because of their relative youth (age 27), most of the alcohol abusers were type 2. As predicted, childhood personality traits of high novelty-seeking, low harm-avoidance, and low reward-dependence strongly increased the risk of alcoholism by age 27. These personality traits distinguished boys who had nearly 20-fold differences in their risk of alcohol abuse: the risk of alcohol abuse varied from four to 75 per cent depending on childhood personality. This prospective study is still ongoing in order to test the prediction that those with onset of alcoholism after age 27 will have anxiety-prone personalities characterized by the opposite profile of traits, particularly high harm-avoidance and high reward-dependence.

CURRENT AND FUTURE DIRECTIONS

The personality traits associated with increased risk of alcoholism are themselves moderately heritable: 40 to 60 per cent (Cloninger, 1987b). They

Table 5. Distinguishing Characteristics of Two Types of Alcoholism

CHARACTERISTIC FEATURES	TYPE OF ALCOHOLISM TYPE 1	TYPE 2
ALCOHOL-RELATED PROBLEMS		
Usual age of onset (years)	After 25	Before 25
Spontaneous alcohol-seeking (inability to abstain)	Infrequent	Frequent
Fighting and arrests when drinking	Infrequent	Frequent
Psychological dependence (loss of control)	Frequent	Infrequent
Guilt and fear (about alcohol dependence)	Frequent	Infrequent
PERSONALITY TRAITS		
Novelty seeking	Low	High
Harm avoidance	High	Low
Reward dependence	High	Low

appear to explain much of the heritable risk of adult alcohol abuse, but pedigree analyses have also identified a possible major gene contributing to risk of type 2 alcoholism (Gilligan *et al.*, 1987). In view of the high heritability of type 2 alcoholism, this pedigree pattern is highly informative for linkage studies: the pedigrees are often loaded with multiple alcoholic relatives in several generations. Consequently, major efforts are now underway to identify specific genetic markers that are associated with increased susceptibility to alcoholism, particularly type 2 alcoholism (Devor & Cloninger, 1989). With the increased knowledge of the human genome, it is now highly probable that a major susceptibility gene could be mapped to specific chromosomal regions if it exists. Possible evidence of linkage has been described for the esterase D (Tanna *et al.*, 1988) and MNS blood group (Hill *et al.*, 1988), but neither conventional levels of statistical significance nor independent confirmation have been demonstrated yet. Other candidate genes for consideration include those related to putative biochemical risk factors, such as variants of adenylate cyclase (Tabakoff *et al.*, 1988), phospholipase D (Mueller *et al.*, 1988), and monoamine oxidase B (von Knorring *et al.*, 1985). Other possible candidate genes may be related to polymorphic variants of enzymes involved in the oxidative or non-oxidative metabolism of alcohol.

The major challenge for future studies will be to explain clinical heterogeneity in terms of the underlying neurochemical and genetic antecedents that influence susceptibility to alcohol abuse (Cloninger, 1987a). Ongoing research is at a watershed because of the confluence of research from nosology and neuropharmacology with genetic research. Old debates about nature versus nurture now appear sterile. It seems clear that genetic factors influence neuroadaptive responses to alcohol intake and other aspects of experience. However, it is now necessary to specify what genes are involved and to describe how they modulate the adaptive responses to alcohol that motivate alcohol abuse.

REFERENCES

Amark, C. (1951). A study in alcoholism. Acta Psychiatr. Neurol. Scand, Suppl. 70.

Bohman, M. (1978). Some genetic aspects of alcoholism and criminality: A population of adoptees. Arch. Gen. Psychiatry 35:269-276.

Bohman, M., Sigvardsson, S., Cloninger, C. R. (1981). Maternal inheritance of alcohol abuse: cross-fostering analysis of adopted women. Arch. Gen. Psychiatry 38:965-969.

Cadoret, R. J., Gath, A. (1978). Inheritance of alcoholism in adoptees. Brit. J. Psychiatry 132:252-258.

Cloninger, C. R. (1987a). Neurogenetic adaptive mechanisms in alcoholism. Science 236:410-416.

Cloninger, C. R. (1987b). A systematic method for clinical description and classification of personality variants: A proposal. Arch. Gen. Psychiatry 44:573-588.

Cloninger, C. R., Bohman, M., Sigvardsson, S. (1981). Inheritance of alcohol abuse: cross-fostering analysis of adopted men. Arch. Gen. Psychiatry 38:861-868.

Cloninger, C. R., Reich, T., Wetzel, R. (1979). Alcoholism and affective disorders: familial associations and genetic models. In: Alcoholism and Affective Disorders, edited by D. Goodwin and C. Erickson. New York, Spectrum, pp. 57-86.

Cloninger, C. R., Reich, T. (1983). Genetic heterogeneity in alcoholism and sociopathy. In: Genetics of Neurological and Psychiatric Disorders, edited by S. S. Kety, L. P. Rowland, R. L. Sidman, S. W. Matthysse. New York, Raven Press, pp. 145-166.

Cloninger, C. R., Sigvardsson, S., Bohman, M. (1988). Childhood personality predicts alcohol abuse in young adults. Alcoholism: Clinical and Experimental Research 12:494-505.

Cotton, N. S. (1979). The familial incidence of alcoholism. J. Stud. Alcohol 40:89-116.

Devor, E. J., Cloninger, C. R. (1989). Genetics of Alcoholism. Ann. Rev. Genetics, in press.

Devor, E. J., Cloninger, C.R. (1989). Genetics of Alcoholism. Ann Rev. Genet. 23:19-36.

Devor, E. J., Reich, T., Cloninger, C. R. (1988). Genetics of alcoholism and related end-organ damage. Sem. Liver Dis. 8:1-11.

Gilligan, S. B., Reich, T., Cloninger, C. R. (1987). Etiologic heterogeneity in alcoholism. Genetic Epidemiology 4:395-414.

Gilligan, S. B., Reich, T., Cloninger, C. R. (1988). Alcohol-related symptoms in heterogeneous families of hospitalized alcoholics. Alcoholism 12:671-678.

Goodwin, D. W. (1976). Is Alcoholism Hereditary? New York, Oxford University Press.

Goodwin, D. W. (1987). Genetic influences in alcoholism. Adv. Intern. Med. 32:283-298.

Goodwin, D. W., Guze, S. B. (1989). Psychiatric Diagnosis. 4th Ed. New York, Oxford University Press.

Goodwin, D. W., Schulsinger, F., Hermansen, L., Guze, S. B., Winokur, G. (1973). Alcohol problems in adoptees raised apart from alcoholic biological parents. Arch. Gen. Psychiatry 28:238-243.

Goodwin, D. W., Schulsinger, F., Knop, J., Mednick, S. A., Guze, S. B. (1977a). Alcoholism and depression in adopted-out daughters of alcoholics. Arch. Gen. Psychiatry 34:751-775.

Goodwin, D. W., Schulsinger, F., Knop, J., Mednick, S. A., Guze, S. B. (1977b). Psychopathology in adopted and non-adopted daughters of alcoholics. Arch. Gen. Psychiatry 34:1005-1009.

Hill, S. Y., Aston, C., Rabin, B. (1988). Suggestive evidence of genetic linkage between alcoholism and the MNS blood group. Alcohol Clin. Exp. Res 12:811-814.

Jellinek, E. M. (1960). Phases of Alcohol Addiction. Q. J. Stud. Alcohol 13:573-584.

Kaij, L. (1960). Alcoholism in Twins: Studies on the Etiology and Sequels of Abuse of Alcohol. Almquist and Wiksell, Stockholm.

Kaij, L. and Dock, J. (1975). Grandsons of alcoholics. Arch. Gen. Psychiatry 32:1379-1381.

Mueller, G. C., Fleming, M. F., LeMahieu, M. A., Lybrand, G.S., Barry, K.J. (1988). Synthesis of phosphatidylethanol — a potential marker for adult males at risk for alcoholism. Proc. Natl. Acad. Sci. USA 85:9778-9782

Partanen, J., Bruun, K., Markkanen, T. (1966). Inheritance of Drinking Behavior. The Finnish Foundation for Alcohol Studies, Helsinki.

Roe, A., Burks, B.(1945). Adult adjustment of foster children of alcoholic and psychotic parentage and the influence of the foster home. No. 3. Memoirs of Section on Alcohol Studies. New Haven, Connecticut, Yale University.

Tabakoff, B., Hoffman, P. L., Lee, J. M., Saito, T., Willard, B., DeLeon-Jones, F. (1988). Differences in platelet enzyme activity between alcoholics and non-alcoholics. New Engl. J. Med. 318:134-139.

Tanna, V. L., Wilson, A. F., Winokur, G., Elston, R. C. (1988). Possible linkage between alcoholism and esterase D. J. Stud. Alcohol 49:472-476.

von Knorring, A.-L., Bohman, M., von Knorring, L., Oreland, L. (1985). Platelet MAO activity as a biological marker in subgroups of alcoholism. Acta Psychiatr. Scand 72:51-58.

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

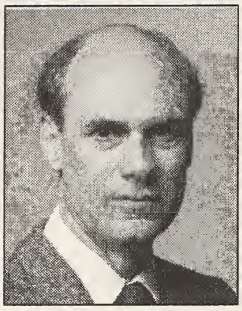
Evolving Concepts of Reinforcement and Drug Self-Administration

ROY A. WISE

DEPARTMENT OF PSYCHOLOGY

CONCORDIA UNIVERSITY

MONTREAL



ROY A. WISE

Roy A. Wise received his PhD in physiological psychology from McGill University in 1968. His interest in addiction derived from a general interest in motivation and a specific interest in hypothalamic electrical stimulation, which he first used to induce ethanol drinking and preference in animals. He has recently been developing two lines of study. One involves the use of microdialysis and *in vivo* voltammetry to measure brain dopamine concentrations in rats during and following drug self-administration. The second involves an assessment of the effects on reward mechanisms of various drug withdrawal syndromes, and the degree to which withdrawal distress adds to the rewarding impact of habit-forming drugs. Recent publications include "Localization of Drug Reward Mechanisms by Intracranial Injections," *Synapse* (in press); and "Reinstatement of heroin self-administration habits: Morphine prompts and naltrexone discourages renewed responding after extinction," *Psychopharmacology* (in press).

The most widely influential theories of drug addiction have held that compulsive drug self-administration is motivated by a need to self-medicate drug-associated withdrawal distress. The influence of this view of addiction is particularly evident from attempts to establish lower animal models of alcoholism; dozens of investigators have based their work on the hypothesis that if they could establish ethanol dependence in their animals, they would thereby guarantee voluntary ethanol consumption. While several methods have been found to induce ethanol dependence, ethanol dependence does not appear to significantly alter the general reluctance of lower animals to ingest intoxicating amounts of ethanol when other sources of fluid and calories are freely available (Falk *et al.*, 1972; Freund, 1973; Goldstein, 1974; Ratcliffe, 1972). Moreover, the methods for producing physical dependence appear to be incompatible with the methods for producing voluntary ethanol intake, at least in rats. To produce dependence, ethanol must be maintained in the body on a continuous basis (Falk *et al.*, 1972; Freund, 1973; Ratcliffe, 1972); however it seems that ethanol must be given on an intermittent basis to produce significant voluntary ethanol intake in rats (Wise, 1975; Amit *et al.*, 1976; Samson and Grant 1990).

Even in the case of opiate addiction, the opiate actions that produce physiological dependence seem independent of those that produce drug-seeking habits. Habit-forming effects of opiates can be demonstrated with opiate doses lower than those needed to produce dependence signs (Deneau *et al.*, 1969; Woods and Schuster, 1971). They can also be demonstrated with low doses of morphine injected directly into the ventral tegmental area of the brain; such injections do not produce physiological dependence signs (Bozarth and Wise, 1984). On the other hand, strong dependence signs can be induced by opiate injections into the periaqueductal gray (Bozarth and Wise, 1984; Wei and Loh, 1976); such injections are not similarly habit-forming (unless much larger doses — large enough to diffuse to the ventral tegmental area — are given).

Thus classical physiological dependence syndromes and drug self-administration habits appear to result from dissociable actions of ethanol and opiates. Moreover, drug self-administration habits can be demonstrated with psychomotor stimulants and nicotine, drugs for which a traditional dependence syndrome is difficult to demonstrate. It appears that physical dependence is not a condition common to the range of classes of abused drugs (Wise, 1987b).

Recent studies of drug self-administration in animals have tended to focus on the habit-forming (rewarding or 'reinforcing') properties of drugs; for the most part, addictive drugs are reinforcing even when the animal does not show classic signs of physical dependence (Collins *et al.*, 1984; Johanson, 1978; Schuster and Thompson, 1969). Behavioral pharmacologists have, in recent years, studied drug self-administration with minimal attention to states of physiological dependence, often documenting the nature of

habit-forming drug effects under conditions designed to minimize the development of dependence. Since drug dependence seems not to be a necessary condition for demonstrating drug reinforcement, drug reinforcement has developed as a topic of study in its own right. Interpretation of drug self-administration patterns has followed lines established in psychological theory.

THEORY OF REINFORCEMENT

Pavlov and Skinner

The term 'reinforcement' was coined by Pavlov (1927)[1] to refer to the strengthening of presumed associations between the central representations of unconditioned and conditioned stimuli. Pavlov's theory held that it was associations between one stimulus and another (rather than between a stimulus and a response) that formed the basis of conditioned reflexes, and Pavlov used the metaphor of reinforcment to indicate the occasional strengthening of association — provided by periodic pairing of the conditioned and unconditioned stimuli — which is necessary if the association is not to be 'extinguished'[2] by repeated presentation of the conditioned stimulus alone.

The term reinforcement was introduced into the literature of instrumental behavior by Thorndike (1933) and Skinner (1933). It was Thorndike who articulated the metaphor that is carried by the term reinforcement; Thorndike (1898) first identified the process as one of the 'stamping in' of stimulus-response associations and subsequently (Thorndike, 1911) associated this process with his 'Law of Effect.' Like Pavlov, Thorndike used the word reinforcement to imply the strengthening not only of mental associations but of connections between neurons or groups of neurons in the brain.

In his initial studies of lever-pressing, Skinner (1933) first worked within the Pavlovian concept of reinforcement, but he later (1937) designated a new category — 'operant' reinforcement — which he distinguished from Pavlovian or 'respondent' reinforcement. Skinner defined operant reinforcement in relation to the development of habits that 'operate' on the environment, distinguishing it from the reinforcement of associations between perceptual events that 'respond' to (are elicited by) the environment. Operant reinforcement was seen by Skinner as involving a completely different learning principle than respondent reinforcement, and he developed a very different experimental paradigm (a variant of the paradigm of Thorndike) for demonstrating the new principle. The principle involved response reinforcement (as contrasted with the Pavlovian principle of stimulus reinforcement or stimulus-stimulus

association); the experimental paradigm was the lever-pressing or key-pecking of rats or pigeons in what came to be known as a 'Skinner box.'

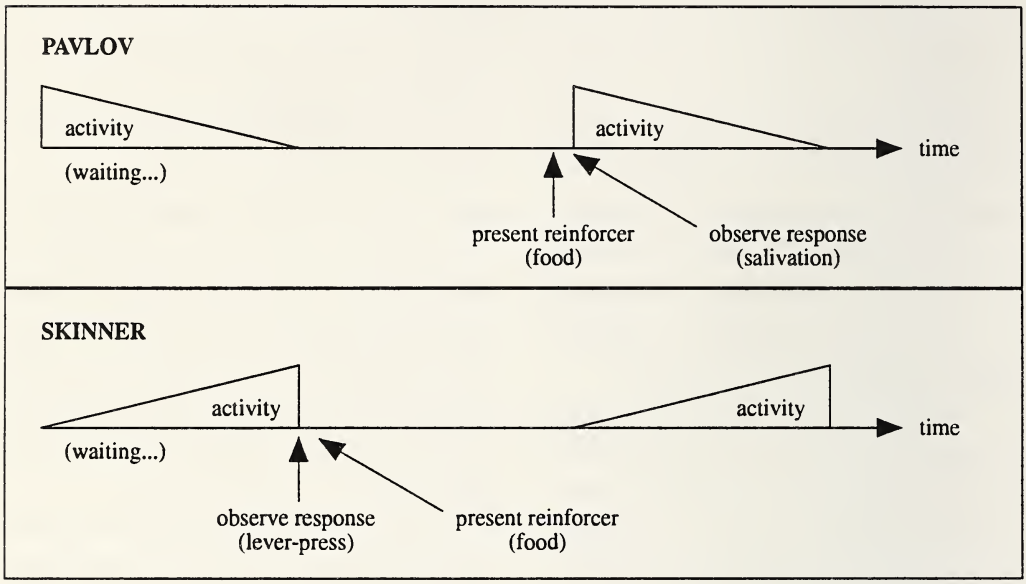
While the study of drug reinforcement by behavioral pharmacologists has developed in the Skinnerian tradition, physiological psychologists have recently developed a Pavlovian paradigm to measure drug reinforcement. Because our understanding of the neurobiology of drug reinforcement now derives from both paradigms, a somewhat detailed discussion of the differences of the two theories and paradigms is important.

In Pavlov's paradigm, the original procedure was to wait until the animal was inactive[3], then to present one or more stimuli, and finally to observe a response. These steps are diagrammed in **Figure 1**; the stimuli are the independent variables and the response is the dependent variable. The response in question is a response to one or both stimuli; the stimuli are seen as the immediate cause of the response. The unconditioned stimulus, often food, is designated the reinforcer.

In Skinner's paradigm, the flow chart runs in reverse. The experimenter waits for the animal to become active and only when the desired act or 'response' occurs is the 'stimulus' — the reinforcer (often food) — presented. **Figure 1** contrasts the Skinnerian paradigm with the Pavlovian. In the Skinnerian paradigm the notion of independent and dependent variables does not readily apply. The stimulus cannot be considered an independent variable as it is not independent of the behavior of the animal (indeed, it is strictly contingent upon the 'response' of the animal); the 'response' cannot be considered an independent variable because it is under the control of the animal and not the experimenter.[4] The 'response' in question is not a response to the manipulated stimulus, as it occurs before the stimulus; the stimulus is contingent upon the response and thus 'caused' by it. Indeed, the response is sometimes said not to be elicited by any identifiable stimulus (Skinner, 1938); rather it is said to be emitted *de novo*, to be favored by reinforcement much as random genetic mutations are favored by natural selection.[5]

The reinforcing effects of drugs have been studied from both the Pavlovian and the Skinnerian point of view. In her classic study of morphine addiction in chimpanzees, Spragg (1940) first noted signs of drug reinforcement in the attraction of her animals to their injection room. The animals began to pull at the leash, attempting to return to the injection room on days when the expected injection was withheld. This behavior was not the result of operant conditioning, as reinforcement was not given following the leash-tugging behavior. Indeed, this response often caused additional omission of reinforcement, as Spragg confirmed her observations on subsequent days. The critical conditioning event here was the Pavlovian association of the injection room with the morphine injections.

Figure 1. Comparison of the Pavlovian and Skinnerian reinforcement paradigms. In the Pavlovian paradigm, the experimenter waited for activity to decrease (for the animal to 'settle down') and then, after a period of inactivity, the experimenter presented the reinforcer. The response of interest to the experimenter followed the reinforcer and was time-locked to it. In the Skinnerian paradigm, the experimenter waited for the activity level to increase to the point that the animal emitted the required response (investigation of the manipulandum) and then, after the required response (and time-locked to it) the reinforcer was presented.



The learning of place-drug associations has subsequently been developed by physiological psychologists as a test of drug reinforcement (see, e.g., Beach, 1957; Rossi & Reid, 1976; Schwartz & Marchok, 1974). Rats are tested for their relative preference for a given portion of an environmental chamber, are then given drug injections while restricted to that portion of the chamber, and are then retested for preference; if their preference for the drug-associated portion of the chamber increases, drug reinforcement is inferred. In the programming of reinforcement, this is a Pavlovian paradigm; the reinforcement is given in association with to-be-conditioned stimuli and is not contingent upon any response of the animal.

Despite the Pavlovian conditioning paradigm, the conditioned response is not, strictly speaking (Wise, 1989a), a simple conditioned response of the sort posited by Pavlov. For Pavlov, the conditioned response was seen as a fraction of the unconditioned response. In the place-preference paradigm, however, it is difficult to identify the conditioned response with any unconditioned response to drug injections. The conditioned response is characterized as approach to and maintenance of contact with the conditioned stimuli of the drug-associated environment. The unconditioned response

cannot be characterized as 'approach' to the unconditioned stimulus, however, as the unconditioned stimulus is a central drug effect that cannot be localized in the environment. It can be said that *forward locomotion* is an unconditioned response to the drug (Wise & Bozarth, 1987), but the conditioned response cannot be characterized as merely forward locomotion. Indeed, if forward locomotion were conditioned to the drug-associated side of the test box, the animal would tend to be active until it left the drug-associated side of the box and to remain inactive — and thus in place — whenever it found itself out of reach of those stimuli. The result would be more time in the neutral side of the apparatus and less time in the drug-associated side, whereas, in fact, we see the opposite. In any case, the paradigm is a Pavlovian paradigm by definition; the animals are given reinforcement in a manner uncorrelated with their own behavior and the behavior of interest follows rather than precedes the reinforcing drug injection.

Thus there are two definitions, two conceptions, and two paradigms of drug reinforcement. Specialists in conditioning theory are not yet agreed as to which conceptions are most adequate. Bindra (1974) has gone so far as to argue that the Skinnerian principle of response reinforcement is invalid, and that all learning results from the Pavlovian form of conditioning. Bindra's argument stems from his insistence on an explanation involving efficient causes — of the sort favored in science since the teachings of David Hume[6] — as are implicit in Pavlov's analysis. Skinner's analysis is in terms of distal or historical causes; Skinner attributes a change in the probability of a given response to the events of reinforcement that occurred in the environmental history of the animal (Skinner, 1985). The implication of Bindra's position is that when the efficient cause of an event is known it is not necessary to appeal to historical causes. Whether or not Bindra is correct, many of the findings of operant psychology are coming to be explained in terms of Pavlovian conditioning and Pavlovian efficient causes (see, e.g., Rescorla, 1987).

The interesting empirical finding for the student of addiction is that the same drugs — and, indeed, the same mechanisms — are implicated in the habit-forming effects of drugs of abuse regardless of which of the two definitions, two conceptions, or two paradigms is used (Wise and Bozarth, 1987; Wise, 1989a).

Mixed Paradigms

The conditioned place preference paradigm, where the drug is given regardless of the animal's behavior, and the lever-press paradigm, where the drug injection is contingent on the animal's responding, represent almost prototypical Pavlovian and Skinnerian situations. A third paradigm that has proven useful in assessment of reinforcement-associated drug effects is

a mixed paradigm where the drug is given in the Pavlovian way but the response is measured in the Skinnerian way. The best studied example involves the brain stimulation reward paradigm (Kornetsky and Esposito, 1979; Kornetsky *et al.*, 1979; Levitt *et al.*, 1977; Wise, 1980; Wise *et al.*, in press). In this paradigm, drug reinforcement is inferred (see, e.g., Broekkamp *et al.*, 1976) from the unconditioned effects of a drug on the operant reinforcement provided by brain stimulation. Here two reinforcers are given. One — the drug — is given in the Pavlovian manner and the other — the brain stimulation — is given in the Skinnerian manner. The unearned effects of the drug are reflected in the rate of responding for the earned brain stimulation. Again, while the paradigm does not fit neatly with either that of Pavlov or that of Skinner, the drugs that facilitate the rewarding effects of brain stimulation are the same drugs that are reinforcing in their own right, as seen in either the conditioned place preference or the lever-pressing situations (Wise and Bozarth, 1987; Wise and Rompré, 1989; Wise, 1989a).

Another mixed paradigm is the reinstatement paradigm (Gerber and Stretch, 1975; Stewart and de Wit, 1987). Here the animal is trained under operant drug reinforcement but tested following response-independent ('priming') drug injections that are given after extinction of the response habit during a period of non-reinforcement. The drug test follows the Pavlovian paradigm, but the animal is prepared under the Skinnerian paradigm. The response that is observed when the training drug is given as the priming stimulus is the lever-pressing response that was established by operant reinforcement. That is to say, the observed response is neither an unconditioned response to the priming drug nor a conditioned response that was established by pairing the priming drug with an unconditioned stimulus for lever-pressing. In this sense the reinstatement paradigm is a hybrid of the operant and Pavlovian paradigms.

NEUROBIOLOGY OF DRUG REINFORCEMENT

Psychomotor stimulants

Drug self-administration studies first suggested involvement of the mesolimbic dopamine system in psychomotor stimulant reinforcement. It is actions in this system that appear to be responsible for those behavioral effects that define these drugs as 'psychomotor stimulants' (Wise and Bozarth 1987). One of the effects of psychomotor stimulants is to increase synaptic dopamine levels, either by causing direct dopamine release from nerve terminals (Axelrod, 1970; Carlsson, 1970) or by blocking the reuptake mechanism that usually terminates synaptic actions of dopamine by removing it from the synapse (Heikkila *et al.*, 1975). Low doses of selective dopamine receptor blockers cause compensatory increases in

lever-pressing for amphetamine (Yokel and Wise, 1975, 1976) and cocaine (de Wit and Wise, 1977; Ettenberg *et al.*, 1982), as do decreases in dose per injection (Yokel and Pickens, 1973, 1974) or increases in work requirements (Pickens and Thompson, 1968). The increases in lever-pressing are presumed to reflect partial blockade of the reinforcing effects of the drugs, such that higher than normal amphetamine or cocaine concentrations are required to produce drug satiety (Wise, 1987a). With higher doses of dopamine blockers, an initial period of accelerated responding is followed by response cessation; this pattern parallels what is seen when neutral saline injections are substituted for reinforcing injections of amphetamine or cocaine (Yokel and Wise, 1975).

Subsequent work showed that dopamine terminals of nucleus accumbens are particularly important for amphetamine and cocaine reinforcement. Dopamine-selective lesions of nucleus accumbens (Roberts *et al.*, 1977; 1980; Lyness *et al.*, 1979) or the ventral tegmental area—site of origin of the dopaminergic fibers projecting to nucleus accumbens—(Roberts and Koob, 1982) cause similar effects to those of neuroleptics. While such lesions damage dopamine fibers to frontal cortex as well as to nucleus accumbens, kainic acid lesions (Zito *et al.*, 1985) and local neuroleptic microinjections (Phillips *et al.*, 1983), which affect only the nucleus accumbens target cells, have similar effects.

Further evidence that nucleus accumbens is important in amphetamine self-administration comes from the finding that rats will work for direct amphetamine injections into this region but not into the overlying ventricle (Hoebel *et al.*, 1983). Similarly, dopamine itself is self-administered into this region (Guerin *et al.*, 1984). Further, amphetamine injections into nucleus accumbens reinstate extinguished drug self-administration habits (Stewart and Vezina, 1988). Cocaine is reported not to be self-administered into this region, but to be self-administered into the dopamine terminal field of the frontal cortex (Goeders & Smith, 1983, 1986).

It is not clear why cocaine should not be self-administered into nucleus accumbens; it certainly is effective in increasing synaptic dopamine levels there (Di Chiara and Imperato, 1988). Since other treatments that elevate nucleus accumbens dopamine levels are reinforcing, it would be expected that local cocaine injections should be reinforcing. Moreover, it is not clear why frontal cortex cocaine injections should be reinforcing, since lesions of frontal cortex do not alter intravenous cocaine self-administration (Martin-Iverson *et al.*, 1986).

Conditioned place preference studies also implicate nucleus accumbens in psychomotor stimulant reinforcement. Dopamine antagonists attenuate amphetamine- (Mackey and van der Kooy, 1985; Mithani *et al.*, 1986; Spyraki *et al.*, 1982a) and cocaine-conditioned (Spyraki *et al.*, 1982b; Mackey and van der Kooy, 1985;

Morency and Beninger, 1986) place preference. The selective dopamine agonists apomorphine (Spyraki *et al.*, 1982a; van der Kooy *et al.*, 1983; Swerdlow *et al.*, 1984), bromocriptine (Hoffman *et al.*, 1988; Morency and Beninger, 1986), and N-0437 (Gilbert *et al.*, 1986) also cause conditioned place preferences. Nucleus accumbens injections of amphetamine induce conditioned place preference, but injections into caudate, amygdala, or frontal cortex do not (Carr and White, 1983, 1986). While nucleus accumbens cocaine is reported not to be self-administered, it is reported to be effective in the conditioned place preference paradigm (Aulisi and Hoebel, 1983).

Finally, nucleus accumbens amphetamine facilitates the reinforcing effects of brain stimulation, increasing self-stimulation rates (Broekkamp *et al.*, 1975) and reducing the amount of stimulation necessary to motivate a given level of performance (Colle and Wise, 1988).

Opiates

The mesolimbic dopamine system is also implicated in opiate reinforcement; opiates appear to have reinforcing actions at the level of the ventral tegmental cell bodies and in the region of the dopamine synapses of nucleus accumbens. The first suggestion that opiates interacted with the mesolimbic dopamine system came from the report of Broekkamp *et al.*, (1976) that morphine facilitates intracranial self-stimulation more effectively when it is injected directly into the region of the dopamine cell bodies. More recent studies have confirmed this finding, suggesting that it is mu and perhaps delta but not kappa receptors that mediate the reward-facilitating effects of ventral tegmental opiates (Jenck *et al.*, 1987).

Subsequent studies have confirmed that morphine is reinforcing in its own right when injected into this region; this is demonstrated by lever-pressing studies (Bozarth and Wise, 1981a; Devine and Wise, 1990; van Ree and de Wied, 1980; Welzl *et al.*, 1989) as well as by conditioned place preference studies (Bozarth and Wise, 1981b; Phillips and LePiane 1980; van der Kooy *et al.*, 1982). The dorsal (Phillips and LePiane, 1980) and the rostral and caudal (Bozarth, 1987) boundaries of the site of action correspond closely to the boundaries of the ventral tegmental dopamine cell group. Ventral tegmental morphine injections also reinstate extinguished lever-pressing that was trained under intravenous heroin or cocaine reinforcement, suggesting not only that the ventral tegmental area plays a role in intravenous opiate reinforcement, but also that the same system plays a role in both opiate and stimulant reinforcement (Stewart, 1984).

While the psychomotor stimulants have their most relevant actions at the dopamine carrier in the membrane of the dopamine nerve terminals themselves[7], the ventral tegmental effects of opiates appear to influence dopamine function indirectly. Early evidence suggested that opioid

receptors were localized on the dopamine cell bodies, but this evidence has now been discounted; intracellular recordings do not show evidence for a direct action of morphine on the ion channels with which opiate receptors are known to be associated (Lacey *et al.*, 1989). Still, local and systemic opiate injections are known to increase dopaminergic cell firing (Gysling and Wang, 1983; Matthews and German, 1984; Ostrowski *et al.*, 1982); this is presumed to account for their ability to increase extracellular dopamine concentration in the nucleus accumbens (Di Chiara and Imperato, 1988; Wood *et al.*, 1980) and overcome (within important limits) the suppressive effects of neuroleptics on brain stimulation reward (Rompré and Wise, 1989a,b). The most likely mechanism would appear to involve disinhibition of the dopamine system, perhaps reflecting the opening of potassium channels (North, 1986) on inhibitory pre-synaptic inputs to the dopamine cells themselves (Wise, 1989b).

In addition to their actions at the level of the dopamine cell bodies, opiates have reinforcing actions at the level of the dopamine nerve terminals of nucleus accumbens. The first evidence for such actions came from studies of intracranial morphine self-administration (Olds, 1982; Goeders *et al.*, 1984) and conditioned place preference (van der Kooy *et al.*, 1982). Recent work has shown that nucleus accumbens opiate injections also facilitate the rewarding effects of brain stimulation (West and Wise, 1988). Earlier reports had suggested that nucleus accumbens opiates did not facilitate brain stimulation reward (Broekkamp *et al.*, 1976) and did not support intracranial self-administration (Bozarth and Wise, 1981a). The failure of the earlier studies may have been due to the fact that opiates injected into the ventral tegmental area have rewarding (Bozarth and Wise, 1981a) and psychomotor stimulant (Kalivas *et al.*, 1983) properties at lower doses than are required with nucleus accumbens injections. Indeed, the psychomotor stimulant properties of opiate injections into nucleus accumbens are generally weak in drug-naive animals. For reasons that are not yet understood, these properties can be enhanced by prior treatment with dopamine neurotoxins, dopamine depleters, and dopamine receptor blockers (Stinus *et al.*, 1985, 1986). Whatever the explanation of differential potency, it appears that reinforcing actions involving at least two sites, the ventral tegmental area and the nucleus accumbens (Britt and Wise, 1983; Vaccarino *et al.*, 1985), contribute to the habit-forming actions of intravenous opiate injections. Interestingly, however, nucleus accumbens injections of morphine do not reinstate lever-pressing habits established under intravenous heroin reinforcement (Stewart and Vezina, 1988); this again suggests that the reinforcing effects of opiates in the ventral tegmental area are more potent than those in the nucleus accumbens.

Three other sites of possible reinforcing actions have been proposed:

the lateral hypothalamus (Olds, 1979; Olds and Williams, 1980; van der Kooy *et al.*, 1982), the periaqueductal gray (van der Kooy *et al.*, 1982), and the hippocampus (Stevens *et al.*, 1991). The first two suggestions are based on reinforcing effects of central morphine injections, and in neither case can diffusion to distal sites of action be ruled out. In the case of the conditioned place preferences that have been reported following periaqueductal gray morphine injections, the effective dose was 40 times higher than the effective ventral tegmental injections; the dose that is effective in the ventral tegmental area is not effective when injected into the periaqueductal gray (Phillips and LePiane, 1980; Bozarth and Wise, 1982). Until it is known, for example, that injections between the periaqueductal gray and the ventral tegmental area are ineffective, the possibility cannot be ruled out that reinforcing periaqueductal gray injections are diffusing to a ventral tegmental or other site of action. Similarly, in the case of lateral hypothalamic injections, doses that are effective when injected into the ventral tegmental area have proven ineffective when injected into the lateral hypothalamus (Bozarth and Wise, 1982). Moreover, complete neurotoxic lesions of the cells of the lateral hypothalamus fail to significantly alter the reinforcing effects of intravenous heroin (Britt and Wise, 1981); thus the intrinsic cells of this region do not appear, on present evidence, to play a critical role in opiate reinforcement. In the case of hippocampal opiate reinforcement, it has recently been reported that morphine and dynorphin A are self-administered into the CA3 region (Stevens *et al.*, 1991). This self-administration is blocked by selective mu but not delta or kappa opioid antagonists. The details of this work have only recently been published and the findings have yet to be explored by other workers.

Other drugs of abuse

Knowledge is minimal regarding the sites of reinforcing action of other drugs of abuse. It has been suggested that psychomotor stimulant actions are common to all drugs of abuse and that this suggests common or at least overlapping mechanisms of action (Wise, 1987b, 1989a; Wise and Bozarth, 1987); direct research on this hypothesis is just beginning. There is evidence that several drugs of abuse in addition to the psychomotor stimulants and the opiates increase dopamine cell firing or turnover. These include ethanol (Di Chiara and Imperato, 1985; Gessa *et al.*, 1985; Mereu, 1987), nicotine (Arqueros *et al.*, 1978; Clarke *et al.*, 1985; Giorguieff-Chesselet *et al.*, 1979; Lichtensteiger *et al.*, 1982), cannabis (Ng Cheon Ton *et al.*, 1988), phencyclidine (Gerhardt *et al.*, 1987), and caffeine (Govoni *et al.*, 1984). Benzodiazepines and barbiturates are exceptions to the rule; if anything they appear to inhibit dopamine function (Cheramy *et al.*, 1977; Corrodi *et al.*, 1966; Weiner *et al.*, 1977; Wood, 1982). Perhaps benzodiazepines

and barbiturates activate the same circuitry at a stage efferent to the dopaminergic contribution; another possibility is that the reinforcing effects of these drugs are mediated by mechanisms independent of those that are involved in psychomotor stimulant and opiate reinforcement.

THEORY OF ADDICTION

Psychological theory distinguishes between positive and negative reinforcement; positive reinforcement results from onset of the pleasures of life and negative reinforcement results from termination of the distresses of life. The habit-forming effects of drugs and brain stimulation that are discussed in the present paper are those associated with positive, not negative, reinforcement. The reinforcing effects of the psychomotor stimulants and of medial forebrain bundle electrical stimulation are not associated with the traditional dependence syndromes and deprivation of these substances does not increase an animal's willingness to work for them; indeed, motivation for drugs and stimulation appears to decrease rather than build up with time since the last reinforcement (Gallistel *et al.*, 1974; Pickens and Harris, 1968; Stewart, 1983).[8] Thus the medial forebrain bundle systems that are associated with drug reinforcement are associated with positive reinforcement, and it is positive reinforcement that offers a common denominator for habit-forming actions of the addictive drugs that act through the medial forebrain bundle mechanism (Wise and Bozarth, 1987; Wise, 1987b, 1989a).

Depressant drugs — opiates, ethanol, barbiturates, benzodiazepines, and others — also produce dependence syndromes, and, in dependent animals, these drugs almost certainly owe some of their rewarding impact to the fact that they terminate withdrawal symptoms. It is important to note, however, that the dependence syndromes associated with different drugs do not offer a common denominator for addiction. The withdrawal syndromes associated with stimulant drugs — cocaine and nicotine, for example — are not only different from those associated with depressants; they are generally opposite in nature. The general state of a subject withdrawn from opiates, ethanol, barbiturates, or benzodiazepines is a state of hyper-excitability (Essig *et al.*, 1969; Fraser *et al.*, 1957; Kalant, 1977; Norton, 1970). In contrast, the general state of a subject withdrawn from cocaine (Jones, 1984) or nicotine (Shiffman, 1979) is a state of decreased arousal. Moreover, even among depressant drugs, the withdrawal syndromes differ to varying degrees, and one depressant has only partial ability to alleviate the withdrawal symptoms associated with another depressant (Fraser, *et al.*, 1957; Kalant, 1977). Thus where a common biological substrate may underlie the positive reinforcing effects of a variety of drugs of abuse — including both stimulants and depressants — different biological substrates underlie the negative

reinforcing effects of these drugs.

Thus, studies of the neurobiology of reinforcement suggest both common and unique factors in the habit-forming effects of drugs of abuse. The common factor is a habit-forming action involving positive reinforcement and associated with the psychomotor stimulant properties typical of abused drugs even when their dominant action is depressant (Wise and Bozarth, 1987). These properties are associated with the pleasure or euphoria that drugs of abuse can produce — at least in drug-sophisticated users (McAuliffe, 1975) — rather than with the withdrawal distress that they can alleviate (Bijerot, 1980; McAuliffe and Gordon, 1974; Wise and Bozarth, 1987; Wise, 1987b, 1989a). The unique factors are habit-forming actions involving negative reinforcement and associated with the several dependence syndromes (Wise, 1987b) and their independent biological mechanisms (Fraser, *et al.*, 1957; Kalant, 1977). Whereas the various forms of stress-relief associated with different classes of abused drugs must contribute to some degree to the continued use of drugs by dependent users, the explanation of the self-administration of drugs prior to (and leading to) development of dependence, and of relapse in successfully detoxified subjects (Stewart *et al.*, 1984), would seem to result primarily from the ability of drugs to serve as positive reinforcers.

FOOTNOTES

[1] The term seems to have first been used in the third lecture on conditioned reflexes, which was given in Madrid in 1903.

[2] Also Pavlov's term.

[3] The modern procedure is to present the stimuli without regard to the animal's behavior. The way it was originally done by Pavlov, it might be argued that the experimenter used operant procedures to condition inactivity.

[4] Skinner himself did not see this problem. He treated the administration of reinforcement as an independent variable, despite the fact that the administration was contingent upon the response of the animal and was thus not strictly at the discretion of the experimenter.

[5] Indeed, while Skinner did not model his reinforcement theory on Darwin's theory of natural selection, he has come to view the two as parallel (Skinner, 1985).

[6] David Hume argued that a cause must satisfy three conditions: it must be contiguous (in both space and time) with its effect, it must precede its effect, and it must have a necessary connection with its effect. This latter idea has evolved into the concept of a "necessary and sufficient condition." While modern conceptions in the field of sub-atomic physics are at variance with Hume's notions of causality, Hume's criteria still designate the kind of cause that is most intuitively satisfying in the behavioral sciences.

[7] In addition to their actions at the dopamine nerve terminals, the psychomotor stimulants appear to affect self-inhibiting feedback at the level of the dopamine cell bodies by increasing dendritic dopamine release and reuptake near cell-body 'autoreceptors.' As will be evident from studies of the opiates, the effects of drugs on dopaminergic self-regulatory feedback can interact with any effects on synaptic function; this source of influence has, until recently (Rompré and Wise, 1989a,b) been largely ignored in attempts to understand dopamine function.

[8] It is true that the probability of responding for intravenous stimulant injections increases during the time in which the last reinforcing injection is being metabolized (Dougherty and Pickens, 1974; Yokel and Pickens, 1973, 1974); it is deprivation beyond this time that is associated with decay of drug motivation (Pickens and Harris 1968; Stewart 1983). Motivation for brain stimulation reinforcement appears to begin to decay the instant the last stimulation is terminated (Gallistel *et al.*, 1974).

REFERENCES

Amit, Z., Sutherland, A., & White, N. (1976). The role of physical dependence in animal models of human alcoholism. *Drug and Alcohol Dependence, 1*, 435-440.

Arqueros, L., Naquira, D., & Zunino, E. (1978). Nicotine-induced release of catecholamines from rat hippocampus and striatum. *Biochemical Pharmacology, 27*, 2667-2674.

Aulisi, E. F., & Hoebel, B. G. (1983). Rewarding effects of amphetamine and cocaine in the nucleus accumbens and block by alpha-flupenthixol. *Society for Neuroscience Abstracts, 9*, 121.

Axelrod, J. (1970). Amphetamine: Metabolism, physiological disposition, and its effects on catecholamine storage. In E. Costa, & S. Garattini (Eds.), *Amphetamines and Related Compounds* (pp. 207-216). New York: Raven Press.

Beach, H. D. (1957). Morphine addiction in rats. *Canadian Journal of Psychology, 11*, 104-112.

Bejerot, N. (1980). Addiction to pleasure: A biological and social-psychological theory of addiction. In D. J. Lettieri, M. Sayers, & H. W. Pearson (Eds.), *Theories on Drug Abuse: Selected Contemporary Perspectives* (pp. 246-255). Rockville, MD: National Institute on Drug Abuse.

Bindra, D. (1974). A motivational view of learning, performance, and behavior modification. *Psychological Review, 81*, 199-213.

Bozarth, M. A. (1987). Neuroanatomical boundaries of the reward-relevant opiate-receptor field in the ventral tegmental area as mapped by the conditioned place preference method in rats. *Brain Research, 414*, 77-84.

Bozarth, M. A., & Wise, R. A. (1981a). Intracranial self-administration of morphine into the ventral tegmental area of rats. *Life Sciences, 28*, 551-555.

Bozarth, M. A., & Wise, R. A. (1981b). Heroin reward is dependent on a dopaminergic substrate. *Life Sciences, 29*, 1881-1886.

Bozarth, M. A., & Wise, R. A. (1982). Localization of the reward-relevant opiate receptors. In L. S. Harris (Ed.), *Problems of Drug Dependence, 1981 (National Institute on Drug Abuse Research Monograph 41)* (pp. 158-164). Washington, DC: U.S. Government Printing Office.

Bozarth, M. A., & Wise, R. A. (1984). Anatomically distinct opiate receptor fields mediate reward and physical dependence. *Science, 224*, 516-517.

Britt, M. D., & Wise, R. A. (1981). Opiate rewarding action: Independence of the cells of the lateral hypothalamus. *Brain Research, 222*, 213-217.

Britt, M. D., & Wise, R. A. (1983). Ventral tegmental site of opiate reward: Antagonism by a hydrophilic opiate receptor blocker. *Brain Research, 258*, 105-108.

Broekkamp, C. L. E., Pijnenburg, A. J. J., Cools, A. R., & Van Rossum, J. M. (1975). The effect of microinjections of amphetamine into the neostriatum and the nucleus accumbens on self-stimulation behavior. *Psychopharmacologia, 42*, 179-183.

Broekkamp, C. L. E., Van den Bogaard, J. H., Heijnen, H. J., Rops, R. H., Cools, A. R., & Van Rossum, J. M. (1976). Separation of inhibiting and stimulating effects of morphine on self-stimulation behavior by intracerebral microinjections. *European Journal of Pharmacology, 36*, 443-446.

Carlsson, A. (1970). Amphetamine and brain catecholamines. In E. Costa, & S. Garattini (Eds.), *Amphetamines and Related Compounds* (pp. 289-300). New York: Raven Press.

Carr, G. D., & White, N. M. (1983). Conditioned place preference from intra-accumbens but not intra-caudate amphetamine injections. *Life Sciences, 33*, 2551-2557.

Carr, G. D., & White, N. M. (1986). Anatomical dissociation of amphetamine's rewarding and aversive effects: An intracranial microinjection study. *Psychopharmacology, 89*, 340-346.

Cheramy, A., Nieoullon, A., & Glowinski, J. (1977). Blockade of the picrotoxin-induced in vivo release of dopamine in the cat caudate nucleus by diazepam. *Life Sciences, 20*, 811-816.

Clarke, P. B. S., Hommer, D. W., Pert, A., & Skirboll, L. R. (1985). Electrophysiological actions of nicotine on substantia nigra single units. *British Journal of Pharmacology, 85*, 827-835.

Colle, L. M., & Wise, R. A. (1988). Effects of nucleus accumbens amphetamine on lateral hypothalamic brain stimulation reward. *Brain Research, 459*, 361-368.

Collins, R. J., Weeks, J. R., Cooper, M. M., Good, P. I., & Russell, R. R. (1984). Prediction of abuse liability of drugs using I.V. self-administration by rats. *Psychopharmacology, 82*, 6-13.

Corrodi, H., Fuxe, K., & Hokfelt, T. (1966). The effects of barbiturates on the activity of the catecholamine neurons in the rat brain. *Journal of Pharmacy and Pharmacology, 18*, 556-558.

de Wit, H., & Wise, R. A. (1977). Blockade of cocaine reinforcement in rats with the dopamine receptor blocker pimozide but not with the noradrenergic blockers phentolamine or phenoxybenzamine. *Canadian Journal of Psychology, 31*, 195-203.

Deneau, G., Yanagita, T., & Seevers, M. H. (1969). Self-administration of psychoactive substances by the monkey: A measure of psychological dependence. *Psychopharmacologia, 16*, 30-48.

Devine, D., & Wise, R. A. (1990) Self-administration of morphine, [D-Ala2,N-Me-Phe4-Gly5-OL]-enkephalin (DAGO), and [D-Pen2,D-Pen5]-enkephalin (DPDPE) into the ventral tegmentum of the rat. *Society for Neuroscience Abstracts, 16*, 930.

Di Chiara, G., & Imperato, A. (1985). Ethanol preferentially stimulates dopamine release in the nucleus accumbens of freely moving rats. *European Journal of Pharmacology, 115*, 131-132.

Di Chiara, G., & Imperato, A. (1988). Drugs of abuse preferentially stimulate dopamine release in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences* (U.S.), *85*, 5274-5278.

Dougherty, J. D., & Pickens, R. (1974). Effects of phenobarbital and SKR 525A on cocaine self-administration in rats. *Drug Addiction, 3*, 135-143.

Essig, C. F., Jones, B. E., & Lam, R. C. (1969). The effect of pentobarbital on alcohol withdrawal in dogs. *Archives of Neurology, 20*, 554-558.

Ettenberg, A., Pettit, H. O., Bloom, F. E., & Koob, G. F. (1982). Heroin and cocaine intravenous self-administration in rats: Mediation by separate neural systems. *Psychopharmacology, 78*, 204-209.

Falk, J. L., Sampson, H. M., & Winger, G. (1972). Behavioural maintenance of high concentrations of blood ethanol and physical dependence in the rat. *Science, 177*, 811-813.

Fraser, H. F., Wikler, A., Isbell, H., & Johnson, N. K. (1957). Partial equivalence of chronic alcohol and barbiturate intoxications. *Quarterly Journal of Studies on Alcohol, 18*, 541-551.

Freund, G. (1973). Alcohol, barbiturate and bromide withdrawal syndrome in mice. *Annals of the New York Academy of Sciences, 215*, 224-234.

Gallistel, C. R., Stellar, J. R., & Bubis, E. (1974). Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *Journal of Comparative and Physiological Psychology, 87*, 848-859.

Gerber, G. J., Stretch, R. (1975) Drug-induced reinstatement of extinguished self-administration behavior in monkeys. *Pharmacology Biochemistry and Behavior, 3*,1055-1061.

Gerhardt, G. A., Pang, K., & Rose, G. M. (1987). In vivo electrochemical demonstration of the presynaptic actions of phencyclidine in rat caudate nucleus. *Journal of Pharmacology and Experimental Therapeutics, 241*, 714-721.

Gessa, G. L., Muntoni, F., Collu, M., Vargiu, L., & Mereu, G. (1985). Low doses of ethanol activate dopaminergic neurons in the ventral tegmental area. *Brain Research, 348*, 201-204.

Gilbert, D. B., Dembski, J. E., Stein, L., & Belluzzi, J. D. (1986). Dopamine and reward: Conditioned place preference induced by dopamine D2 receptor agonist. *Society for Neuroscience Abstracts, 12*, 938.

Giorguieff-Chesselet, M. F., Kemel, M. L., Wandscheer, D., & Glowinski, J. (1979). Regulation of dopamine release by presynaptic nicotinic receptors in rat striatal slices: effect of nicotine in low concentration. *Life Sciences, 25*, 1257-1262.

Goeders, N. E., Lane, J. D., & Smith, J. E. (1984). Self-administration of methionine enkephalin into the nucleus accumbens. *Pharmacology Biochemistry and Behavior, 20*, 451-455.

Goeders, N. E., & Smith, J. E. (1983). Cortical dopaminergic involvement in cocaine reinforcement. *Science, 221*, 773-775.

Goeders, N. E., & Smith, J. E. (1986). Reinforcing properties of cocaine in the medial prefrontal cortex: Primary action on presynaptic dopaminergic terminals. *Pharmacology Biochemistry and Behavior, 25*, 191-199.

Goldstein, D. B. (1974). Rates of onset and decay of alcohol physical dependence in mice. *Journal of Pharmacology and Experimental Therapeutics, 140*, 377-383.

Govoni, S., Petkov, V. V., Montefusco, O., Missale, C., Battaini, F., Spano, P. F., & Trabucchi, M. (1984). Differential effects of caffeine on dihydroxyphenylacetic concentrations in various rat brain regions. *Journal of Pharmacy and Pharmacology, 36*, 458-460.

Guerin, G. F., Goeders, N. E., Dworkin, S. I., & Smith, J. E. (1984). Intracranial self-administration of dopamine into the nucleus accumbens. *Society for Neuroscience Abstracts, 10*, 1072.

Gysling, K., & Wang, R. Y. (1983). Morphine-induced activation of A10 dopamine neurons in the rat. *Brain Research, 277*, 119-127.

Heikkila, R. E., Orlansky, H., & Cohen, G. (1975). Studies on the distinction between uptake inhibition and release of (3 H)dopamine in rat brain tissue slices. *Biochemical Pharmacology, 24*, 847-852.

Hoebel, B. G., Monaco, A. P., Hernandez, L., Aulisi, E. F., Stanley, B. G., & Lenard, L. (1983). Self-injection of amphetamine directly into the brain. *Psychopharmacology, 81*, 158-163.

Hoffman, D. C., Dickson, P. R., & Beninger, R. J. (1988). The dopamine D2 receptor agonists, quinpirole and bomocriptine produce conditioned place preferences. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 12*, 315-322.

Jenck, F., Gratton, A., & Wise, R. A. (1987) Opioid receptor subtypes associated with ventral tegmental facilitation of lateral hypothalamic brain stimulation reward. *Brain Research, 423*, 34-38.

Johanson, C. E. (1978). Drugs as reinforcers. In D. E. Blackman, & D. J. Sanger (Eds.), *Contemporary Research in Behavioral Pharmacology* (pp. 325-390). New York: Plenum Press.

Jones, R. T. (1984). The pharmacology of cocaine. In J. Grabowski (Ed.), *Cocaine: Pharmacology, Effects, and Treatment of Abuse* (National Institute on Drug Abuse Research Monograph 50) (pp. 34-53). Washington, DC: U.S. Government Printing Office.

Kalant, H. (1977). Comparative aspects of tolerance to, and dependence on, alcohol, barbiturates, and opiates. In M. M. Gross (Ed.), *Alcohol Intoxication and Withdrawal III* (pp. 169-186). New York: Plenum.

Kalivas, P. W., Widerlov, E., Stanley, D., Breese, G., & Prange, A. J. (1983). Enkephalin action on the mesolimbic system: A dopamine-dependent and a dopamine-independent increase in locomotor activity. *Journal of Pharmacology and Experimental Therapeutics, 227*, 229-237.

Kornetsky, C., & Esposito, R. U. (1979). Euphorigenic drugs: Effects on the reward pathways of the brain. *Federation Proceedings, 38*, 2473-2476.

Kornetsky, C., Esposito, R. U., McLean, S., & Jacobson, J. O. (1979). Intracranial self-stimulation thresholds: A model for the hedonic effects of drugs of abuse. *Archives of General Psychiatry, 36*, 289-292.

Lacey, M. G., Mercuri, N. B., & North, R. A. (1989). Two cell types in rat substantia nigra zona compacta distinguished by membrane properties and the actions of dopamine and opioids. *Journal of Neuroscience, 9*, 1233-1241.

Levitt, R. A., Baltzer, J. H., Evers, T. M., Stilwell, D. J., & Furby, J. E. (1977). Morphine and shuttle-box self-stimulation in the rat: a model for euphoria. *Psychopharmacologia, 54*, 307-311.

Lichtensteiger, W., Hefti, F., Felix, D., Huwyler, T., Melamed, E., & Schlumpf, M. (1982). Stimulation of nigrostriatal dopamine neurones by nicotine. *Neuropharmacology, 21*, 963-968.

Lyness, W. H., Friedle, N. M., & Moore, K. E. (1979). Destruction of dopaminergic nerve terminals in nucleus accumbens: effect on d-amphetamine self-administration. *Pharmacology Biochemistry and Behavior, 11*, 553-556.

Mackey, W. B., & van der Kooy, D. (1985). Neuroleptics block the positive reinforcing effects of amphetamine but not of morphine as measured by place conditioning. *Pharmacology Biochemistry and Behavior, 22*, 101-106.

Martin-Iverson, M. T., Szostak, C., & Fibiger, H. C. (1986). 6-Hydroxydopamine lesions of the medial prefrontal cortex fail to influence intravenous self-administration of cocaine. *Psychopharmacology, 88*, 310-314.

Matthews, R. T., & German, D. C. (1984). Electrophysiological evidence for excitation of rat ventral tegmental area dopaminergic neurons by morphine. *Neuroscience, 11*, 617-626.

McAuliffe, W. E. (1975). A second look at first effects: The subjective effects of opiates on nonaddicts. *Journal of Drug Issues, 5*, 369-399.

McAuliffe, W. E., & Gordon, R. A. (1974). A test of Lindesmith's theory of addiction: The frequency of euphoria among long-term addicts. *American Journal of Sociology, 79*, 795-840.

Mereu, G., Yoon, K.-W. P., Boi, V., Gessa, G. L., Naes, L., & Westfall, T. C. (1987a). Preferential stimulation of ventral tegmental area dopaminergic neurons by nicotine. *European Journal of Pharmacology, 141*, 395-400.

Mithani, S., Martin-Iverson, M. T., Phillips, A. G., & Fibiger, H. C. (1986). The effects of haloperidol on amphetamine- and methylphenidate-induced conditioned place preferences and locomotor activity. *Psychopharmacology, 90*, 247-252.

Morency, M. A., & Beninger, R. J. (1986). Dopaminergic substrates of cocaine-induced place conditioning. *Brain Research, 399*, 33-41.

Ng Cheong Ton, J. M., Gerhardt, G. A., Friedemann, M., Etgen, A., Rose, G. M., Sharpless, N. S., & Gardner, E. L. (1988). The effects of Δ9-tetrahydrocannabinol on potassium-evoked release of dopamine in the rat caudate nucleus: An in vivo electrochemical and in vivo dialysis study. *Brain Research, 451*, 59-68.

North, R. A. (1986). Opioid receptor types and membrane ion channels. *Trends in NeuroScience, 9*, 114-117.

Norton, P. R. E. (1970). The effects of drugs on barbiturate withdrawal convulsions in the rat. *Journal of Pharmacy and Pharmacology, 22*, 763- 766.

Olds, M. E. (1979). Hypothalamic substrate for the positive reinforcing properties of morphine in the rat. *Brain Research, 168*, 351-360.

Olds, M. E. (1982). Reinforcing effects of morphine in the nucleus accumbens. *Brain Research, 237*, 429-440.

Olds, M. E., & Williams, K. N. (1980). Self-administration of D-ala2-met- enkephalinamide at hypothalamic self-stimulation sites. *Brain Research, 194*, 155-170.

Ostrowski, N. L., Hatfield, C. B., & Caggiula, A. R. (1982). The effects of low doses of morphine on the activity of dopamine containing cells and on behavior. *Life Sciences, 31*, 2347-2350.

Pavlov, I. P. (1927). *Conditioned Reflexes*. Oxford: Oxford University Press.

Phillips, A. G., Broekkamp, C. L. E., & Fibiger, H. C. (1983). Strategies for studying the neurochemical substrates of drug reinforcement in rodents. *Progress in Neuro-psychopharmacology and Biological Psychiatry, 7*, 585-590.

Phillips, A. G., & LePiane, F. G. (1980). Reinforcing effects of morphine microinjection into the ventral tegmental area. *Pharmacology Biochemistry and Behavior, 12*, 965-968.

Pickens, R., & Harris, W. C. (1968). Self-administration of d-amphetamine by rats. *Psychopharmacologia, 12*, 158-163.

Pickens, R., & Thompson, T. (1968). Cocaine-reinforced behavior in rats: Effects of reinforcement magnitude and fixed-ratio size. *Journal of Pharmacology and Experimental Therapeutics, 161*, 122-129.

Ratcliffe, F. (1972). Ethanol dependence in the rat: Its production and characteristics. *Archives Internationales de Pharmacodynamie et de Therapie, 196*, 146-156.

Rescorla, R. A. (1987). A Pavlovian analysis of goal-directed behavior. *American Psychologist, 42*, 119-129.

Roberts, D. C. S., Corcoran, M. E., & Fibiger, H. C. (1977). On the role of ascending catecholaminergic systems in intravenous self-administration of cocaine. *Pharmacology, Biochemistry and Behavior, 6*, 615-620.

Roberts, D. C. S., & Koob, G. (1982). Disruption of cocaine self-administration following 6-hydroxydopamine lesions of the ventral tegmental area in rats. *Pharmacology, Biochemistry and Behavior, 17*, 901-904.

Roberts, D. C. S., Koob, G. F., Klonoff, P., & Fibiger, H. C. (1980). Extinction and recovery of cocaine self-administration following 6-OHDA lesions of the nucleus accumbens. *Pharmacology, Biochemistry and Behavior, 12*, 781-787.

Rompré, P.-P., & Wise, R. A. (1989a). Opioid-neuroleptic interaction in brainstem self-stimulation. *Brain Research, 477*, 144-151.

Rompré, P.-P., & Wise, R. A. (1989b). Behavioral evidence for midbrain dopamine depolarization inactivation. *Brain Research, 477*, 152-156.

Rossi, N. A., & Reid, L. D. (1976). Affective states associated with morphine injections. *Physiological Psychology, 4*, 269-274.

Samson, H. H. & Grant, K. A. (1990). Some implications of animal alcohol self-administration studies for human alcohol problems. *Drug and Alcohol Dependence, 25*, 141-144.

Schuster, C. R., & Thompson, T. (1969). Self-administration of and behavioral dependence on drugs. *Annual Review of Pharmacology, 9*, 483-502.

Schwartz, A. S., & Marchok, P. L. (1974). Depression of morphine-seeking behavior by dopamine inhibition. *Nature, 248*, 257-258.

Shiffman, S. (1979). The tobacco withdrawal syndrome. In N. A. Krasnegor (Ed.), *Cigarette Smoking as a Dependence Process (National Institute on Drug Abuse Research Monograph)* (pp. 158-184). Washington, DC: U.S. Government Printing Office.

Skinner, B. F. (1933). On the rate of extinction of a conditioned reflex. *Journal of General Psychology, 8*, 114-129.

Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller. *Journal of General Psychology, 16*, 272-279.

Skinner, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1985). Cognitive science and behaviourism. *British Journal of Psychology, 76*, 291-301.

Spragg, S. D. S. (1940). Morphine addiction in chimpanzees. *Comparative Psychology Monographs, 15*, 1-132.

Spyraki, C., Fibiger, H. C., & Phillips, A. G. (1982a). Dopaminergic substrates of amphetamine-induced place preference conditioning. *Brain Research, 253*, 185-193.

Spyraki, C., Fibiger, H. C., & Phillips, A. G. (1982b). Cocaine-induced place preference conditioning: lack of effects of neuroleptics and 6- hydroxydopamine lesions. *Brain Research, 253*, 195-203.

Stevens, K. E., Shiotsu, G., & Stein, L. (1991) Hippocampal mreceptors mediate opioid reinforcement in the CA3 region. *Brain Research, 545*, 8-16.

Stewart, J. (1983). Conditioned and unconditioned drug effects in relapse to opiate and stimulant drug self-administration. *Progress in Neuro-psychopharmacology & Biological Psychiatry, 7*, 591-597.

Stewart, J. (1984) Reinstatement of heroin and cocaine self-administration behavior in the rat by intracerebral application of morphine in the ventral tegmental area. *Pharmacology, Biochemistry and Behavior, 20*, 917-923.

Stewart, J. & de Wit, H. (1987). Reinstatement of drug-taking behavior as a method of assessing incentive motivational properties of drugs. In M. A. Bozarth (Ed.), *Methods of Assessing the Reinforcing Properties of Abused Drugs* (pp. 211-227). New York: Springer Verlag.

Stewart, J., de Wit, H., & Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in the self- administration of opiates and stimulants. *Psychological Review, 91*, 251-268.

Stewart, J., & Vezina, P. (1988) A comparison of the effects of intra-accumbens injections of amphetamine and morphine on reinstatement of heroin intravenous self-administration behavior. *Brain Research, 457*, 287-294.

Stinus, L., Nadaud, D., Jauregui, J., & Kelley, A. E. (1986). Chronic treatment with five different neuroleptics elicits behavioral supersensitivity to opiate infusion into the nucleus accumbens. *Biological Psychiatry, 21*, 34-48.

Stinus, L., Winnock, M., & Kelley, A. E. (1985). Chronic neuroleptic treatment and mesolimbic dopamine denervation induce behavioural supersensitivity to opiates. *Psychopharmacology, 85*, 323-328.

Swerdlow, N. R., Swanson, L. W., & Koob, G. F. (1984). Electrolytic lesions of the substantia innominata and lateral preoptic area attenuate the 'supersensitive' locomotor response to apomorphine resulting from denervation of the nucleus accumbens. *Brain Research, 306*, 141-148.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs, 8*, 1-109.

Thorndike, E. L. (1911). *Animal intelligence*. Macmillan, New York.

Thorndike, E. L. (1933). A theory of the action of the after-effects of a connection upon it. *Psychological Review, 40*, 434-439.

Vaccarino, F. J., Bloom, F. E., & Koob, G. F. (1985). Blockade of nucleus accumbens opiate receptors attenuates intravenous heroin reward in the rat. *Psychopharmacology, 86*, 37-42.

van der Kooy, D., Mucha, R. F., O'Shaughnessy, M., & Bucenieks, P. (1982). Reinforcing effects of brain microinjections of morphine revealed by conditioned place preference. *Brain Research, 243*, 107-117.

van der Kooy, D., Swerdlow, N. R., & Koob, G. F. (1983). Paradoxical reinforcing properties of apomorphine: Effects of nucleus accumbens and area postrema lesions. *Brain Research, 259*, 111-118.

van Ree, J. M., & de Wied, D. (1980). Involvement of neurohypophyseal peptides in drug-mediated adaptive responses. *Pharmacology, Biochemistry and Behavior, 13* (Suppl. 1), 257-263.

Wei, E. T., & Loh, H. H. (1976). Physical dependence on opiate-like peptides. *Science, 193*, 1262-1263.

Weiner, W. J., Goetz, C., Nausieda, P. A., & Klawans, H. L. (1977). Clonazepam and dopamine-related stereotyped behavior. *Life Sciences, 21*, 901-906.

Welzl, H., Kuhn, G., & Huston, J. P. (1989) Self-administration of small amounts of morphine through glass micropipettes into the ventral tegmental area of the rat. *Neuropharmacology, 28*, 1017-1023.

West, T. E. G., & Wise, R. A. (1988). Nucleus accumbens opioids facilitate brain stimulation reward. *Society for Neuroscience Abstracts, 14*, 1102.

Wise, R. A. (1975). Maximization of ethanol intake in the rat. In M. M. Gross (Ed.), *Alcohol Intoxication and Withdrawal: Experimental Studies II* (pp. 279-294). New York: Plenum.

Wise, R. A. (1980). Action of drugs of abuse on brain reward systems. *Pharmacology, Biochemistry and Behavior, 13* (Suppl. 1), 213-223.

Wise, R. A. (1987a). Intravenous drug self-administration: A special case of positive reinforcement. In M. A. Bozarth (Ed.), *Methods of Assessing the Reinforcing Properties of Abused Drugs* (pp. 117-141). New York: Springer Verlag.

Wise, R. A. (1987b). The role of reward pathways in the development of drug dependence. *Pharmacology and Therapeutics, 35*, 227-263.

Wise, R. A. (1989a). The brain and reward. In J. M. Liebman, & S. J. Cooper (Eds.), *The Neuropharmacological Basis of Reward* (pp. 377-424). Oxford: Oxford University Press.

Wise, R. A. (1989b). Opiate reward: Sites and substrates. *Neuroscience and Biobehavioral Reviews, 13*, 129-133

Wise, R. A., Bauco, P., Carlezon, W. A., Jr., and Trojniar, W. (in press). Self-stimulation and drug reward mechanisms. *Annals of the New York Academy of Sciences.*

Wise, R. A., & Bozarth, M. A. (1987). A psychomotor stimulant theory of addiction. *Psychological Review, 94*, 469-492.

Wise, R. A., & Rompré, P.-P. (1989). Brain dopamine and reward. *Annual Review of Psychology, 40*, 191-225.

Wood, P. L. (1982). Actions of GABAergic agents on dopamine metabolism in the nigrostriatal pathway of the rat. *Journal of Pharmacology and Experimental Therapeutics*, 222, 674-679.

Wood, P. L., Stotland, M., Richard, J. W., & Rackham, A. (1980). Actions of mu, kappa, sigma, delta and agonist/antagonist opiates on striatal dopaminergic function. *Journal of Pharmacology and Experimental Therapeutics*, *215*, 697-703.

Woods, J. H., & Schuster, C. R. (1971). Opiates as reinforcing stimuli. In Thompson & Pickens (Eds.), *Stimulus Properties of Drugs* (pp. 163-175). New York: Appleton-Century-Crofts.

Yokel, R. A., & Pickens, R. (1973). Self-administration of optical isomers of amphetamine and methylamphetamine by rats. *Journal of Pharmacology and Experimental Therapeutics, 187*, 27-33.

Yokel, R. A., & Pickens, R. (1974). Drug level of d- and l-amphetamine during intravenous self-administration. *Psychopharmacologia*, *34*, 255-264.

Yokel, R. A., & Wise, R. A. (1975). Increased lever pressing for amphetamine after pimozide in rats: Implications for a dopamine theory of reward. *Science, 187*, 547-549.

Yokel, R. A., & Wise, R. A. (1976). Attenuation of intravenous amphetamine reinforcement by central dopamine blockade in rats. *Psychopharmacology, 48*, 311-318.

Zito, K. A., Vickers, G., & Roberts, D. C. S. (1985). Disruption of cocaine and heroin self-administration following kainic acid lesions of the nucleus accumbens. *Pharmacology, Biochemistry and Behavior, 25*, 1029-1036.

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

The Evolution of Treatment for Alcohol Problems From the 1940s to the 1980s

WILLIAM R. MILLER

DEPARTMENT OF PSYCHOLOGY

UNIVERSITY OF NEW MEXICO

ALBUQUERQUE, NEW MEXICO

WILLIAM R. MILLER

William R. Miller first became interested in addiction research during a 1973 summer internship on the alcohol treatment unit at the Wood Veterans Administration Medical Centre in Milwaukee. After completing his PhD, he interned at another VA hospital and then accepted a faculty post at the University of New Mexico, where he is now Professor of Psychology and Psychiatry. Dr. Miller developed and was founding director of the Research Division of UNM's Centre on Alcoholism, Substance Abuse and Addiction. His research — now encompassing 17 books and more than 100 articles and chapters — focuses on evaluation of the relative effectiveness of treatments for alcoholism. His works include: *Handbook of Alcoholism Treatment Approaches* (1989); *How to Control Your Drinking* (1976); *Motivational Interviewing* (1991); and several books on the integration of spiritual and behavioral approaches to change.

The year was 1949. The North American continent was recovering from World War II, and was about to embark on the Korean War. Harry Truman was serving his second term as President of the United States. In Canada, the Liberals were in power, and Louis St. Laurent had just been elected Prime Minister. The Addiction Research Foundation was a neonate.

It was a transitional time in the alcohol field. There were still large and vocal groups of 'Drys' to be reckoned with, who favored prohibition and regarded any drinking as pathological. The liquor industry was also a growing political force, and would soon succeed in fixing a U.S. alcohol tax rate that would not be increased for more than three decades. Alcoholics Anonymous had been in existence for 14 years, and public opinion about alcoholism was shifting away from a moral model in favor of more humane treatment. Drunk driving was a growing concern. The *Quarterly Journal of Studies on Alcohol* was issuing Volume 10.

What was treatment like in 1949, and how has the treatment of alcohol problems changed in the ensuing 40 years? This is the challenging question that was posed to me when I was invited to present this lecture. I propose to contrast present treatment practices with those in place in 1949, using four domains for organizational purposes: **1.** conceptual *models* and assumptions underlying treatment efforts; **2.** treatment *systems* and providers; **3.** alternative treatment *modalities*; and **4.** program evaluation *research*. Audaciously armed with the limited longitudinal perspective provided by two data points, I will then attempt to project what changes may lie ahead for the treatment of alcohol problems in the 21st century.

CONCEPTUAL MODELS

The journals of the 1940s reflect considerable ambivalence about what to call alcohol problems. 'Alcoholism' and 'alcoholic' were widely used terms, but were given many different definitions. The most common use of the term 'alcoholism' seems to have been as a generic descriptor, referring to a wide range of alcohol-related problems. Some preferred the term 'problem drinker' as the generic term, with alcoholism representing a subset of problem drinkers.

It was taken for granted that there are several different *subtypes* within this broad category. There was a call to differentiate between true alcohol addiction and the ordinary adverse consequences of chronic intoxication (Bacon, 1949). In a paper that captures well the tenor of treatment in the 1940s, Bowman and Jellinek (1941) wrote, "The question of the treatment of alcohol addiction has been obfuscated to no small degree by failure to differentiate between addiction and symptomatic inebriety" (p. 144). We have here a clear parallel to the current interest in alcohol dependence, and

the attempt to differentiate it from alcohol abuse. The groundwork for typologies was already in place, and with it the conceptual basis for client-treatment matching. In the same paper, Bowman and Jellinek assert that "a definite effort has to be made to establish criteria for the suitability of any given method to a given patient; the criteria so far established are definitely superficial" (p. 170). Voegtlin and Lemere (1942) similarly decried "our inability to properly select which patient will be cured by any particular therapy" (p. 797).

At the same time, the seeds of a unitary model of alcoholism were also being sown. The original 'big book' of Alcoholics Anonymous (1939) set forth the view that alcoholics differ in a quantum and qualitative fashion from nonalcoholics. Jellinek published results of the AA Grapevine survey, and interpreted these data as supporting an orderly progression of symptoms through several phases of alcoholism (Jellinek, 1946). This implies the progression of a *single* condition, a view that Jellinek would later (1960) reject in favor of multiple subtypes. Another, perhaps inadvertent step toward conceptualizing alcoholism as a unitary condition arose with the publication of the first screening instrument designed to make a binary discrimination between 'alcoholics' and 'nonalcoholics.' Manson (1948) developed this first empirically-derived alcoholism scale by administering 470 personality inventory items to two groups: alcoholics (defined by institutionalization or membership in AA) and nonalcoholics (the relatives and friends of alcoholics — probably a poor sample in which to assume the absence of alcohol problems). The psychometric limitations of this instrument notwithstanding (Miller, 1976), it set a conceptual precedent: that one could make a binary distinction between those having or not having alcoholism.

There were, then, some attempts to improve diagnostic sophistication beyond the 'laundry list' approach, which simply enumerates the adverse consequences of frequent intoxication (e.g., Huss, 1849; Rush, 1943). In the 1940s, as now, there existed simultaneously at least three different conceptual approaches. One implied a *continuum* of impairment, culminating with the more severe problems of addiction. Another posited *subtypes* of alcohol problems. The third embraced a *binary* nosology, recognizing only the discrete presence or absence of a unitary condition.

The treatment of alcohol problems was still a relatively new endeavor. The term 'cure' was giving way to the more holistic concept of 'rehabilitation' (Blumestone, 1944). It was clearly recognized that the field possessed no single outstandingly effective treatment method, and that all available approaches had their problems and limitations (Bowman & Jellinek, 1941, p. 45).

There was also, in the 1940s, a striking and consistent recognition of the importance of prevention as a larger

context within which to understand treatment. The first sentence of the first article in the first issue of the first volume of the *Quarterly Journal of Studies on Alcohol* reads, "The continual production of new alcoholics is one of the most serious of all the problems of alcohol" (Henderson, 1940, p. 1). This awareness of environmental impact extended to what might now be called relapse prevention. The overriding importance of the post-treatment environment (cf. Finney, Moos, & Mewborn, 1980) was recognized by Bowman and Jellinek (1941): "One of the immediate difficulties is that part of the alcohol addict's trouble is his environment and a doctor can do very little to alter that. All that can usually be done is to help the patient to live in his given environment" (p. 146). They conclude, "On the whole, we feel that the practical possibilities of prevention are greater than those of cure" (p. 170).

The spectrum of severity of alcohol problems, and the interface of prevention and treatment, were raising questions about the appropriate goals of treatment. Voegtlin and Lemere (1942) pointed out the hypocrisy "seen in the almost universal presentation of the dictum that the goal of therapy must be total and permanent abstinence, to be followed by a presentation of results which state that a certain percentage of patients are 'improved'. If such is possible, the goal of psychotherapy should be that of controlled drinking rather than total abstinence, for the former intent would certainly remove one of the major objections of the patient to accepting therapy" (p. 796). The same division of favorable outcomes into 'abstinent' and 'improved' has been noted by more recent reviewers (e.g. Emrick, 1974), and the related issue of whether moderation is an acceptable treatment goal continues to vex and perplex the field (Heather & Robertson, 1983).

Another pressing conceptual question for treatment in the 1940s was whether alcohol problems were symptomatic of and secondary to other underlying psychological disorders. Some professionals attributed treatment failures "to too much emphasis on the overcoming of the drinking habit, instead of trying to solve the personality difficulties, and to the lack of sufficient understanding of the psychological mechanisms in habit formation" (Bowman & Jellinek, 1941, p. 157). At the same time, Voegtlin and Lemere (1942) recognized that psychological difficulties may be secondary to alcohol problems: "Psychotherapists are united in condemning any form of therapy that is not directed toward the readjustment of the underlying personality deficit.... However it would seem that such a concept of treatment neglects entirely the possibility of spontaneous readjustment of the personality disorder if under certain circumstances, the patient is held abstinent" (pp. 797-798).

In sum, the 1940s, like the 1980s, were a time of conceptual confusion in the alcohol field. Binary, continuum, and typology approaches existed side by side,

all being used as diagnostic models. Severity of addiction was proposed as an important single dimension for assessment, by which patients might be matched to treatments. Dogma and controversy surrounded the issue of controlled drinking. Verbal recognition of the need for client-treatment matching was linked to the acknowledgment that valid empirical criteria for such matching were lacking. There was confusion about whether and when alcohol problems were primary or secondary to other psychological problems. All of these remain, in the 1980s, as unresolved conceptual issues. Many quotations on these issues in the 1940s could have been written last month.

TREATMENT SYSTEMS

The 1940s were also a time of debate regarding the most appropriate setting in which to treat alcohol problems. Bowman and Jellinek (1941) described differences of opinion "around the question of whether the patient should be treated in his own home or in institutions, and around the type of institution" (p. 145). The feasibility of outpatient detoxification had been described in the early 1930s (Rees, 1932). But then, as now, professional opinion seemed to carry more weight than did research data. Bowman and Jellinek described a colleague who "expressed a preference for treatment in closed institutions, since his impression was that success was no greater in the open ones. Although his statistics actually showed a considerably higher success rate in the open institutions, he explained this away" (p. 147).

Bowman and Jellinek also recognized that private institutions were likely to attract good prognosis patients — the type of person "who is not a true addict, and in whom the problem, therefore, is a comparatively simple one" (p. 169). One is reminded of the recently publicized admission, to a private alcoholism hospital, of Mrs. Kitty Dukakis, who was described in the national press as having only occasional episodes of excessive drinking without significant impairment, during a few months following the lost U.S. presidential election.

What *has* clearly changed since the 1940s, however, is the *availability* of such institutions. A 1944 article estimated that there were only 50 specialized alcoholism treatment units in the United States (Corwin & Cunningham, 1944). The first specialized programs within the Veterans Administration hospitals appeared in the late 1940s (Kersten, 1949). The typical course of stay in specialized hospital programs was described as 14 days, with outpatient aftercare, a norm toward which we seem to be returning. The total cost of this course of treatment was between $175 and $350 (Corwin & Cunningham, 1944), a norm toward which we do not seem to be returning.

The dramatic expansion of alcoholism treatment facilities in the 1970s

and 1980s can be understood as a response to a major change in social policy: the decision that treatment for alcoholism is reimbursable by third-party payment under public and private health plans. This was not so in the 1940s, and a generation of humanitarian advocates persisted in their efforts to have alcoholism recognized as a condition deserving treatment and public support. They succeeded beyond their wildest dreams. These humanitarians of the 1940s and 1950s could not, I expect, have envisioned the insatiable economic monster they would unleash in a for-profit health care marketplace. From the perspective of the 1940s, the U.S. treatment atmosphere of the 1980s was simply unimaginable, with its advertising and cutthroat competition of programs, and the use of coercion and ever-broadening definitions of 'alcoholism' to fill empty hospital beds. Here is a clear difference from the 1940s, when more often advocates pleaded with hospitals to accept an alcoholic. Meanwhile, in the late 1980s, we remain without adequate resources or approaches for treating homeless alcoholics, a group for whom Straus was expressing concern in 1946.

Even in the absence of systemic funding for treatment, the 1940s witnessed a broadening base of professionals and agencies concerned with the treatment of alcohol problems. The interest of medicine in general, and private hospitals in particular, was increasing. Psychotherapists were actively involved in treating alcoholics (Voegtlin & Lemere, 1942). The lay fellowship approach of Alcoholics Anonymous was growing, although members and groups at that time apparently recommended hospital treatment as a precursor to AA involvement (Bowman & Jellinek, 1941). Treatment was handled largely by professionals; it was not yet normative to employ recovering paraprofessional counselors as primary treatment agents. Organized religious groups and clergy were becoming actively involved in treatment (Rice, 1943). The matching of patients to treatments was rare; most programs offered their ministrations to all applicants. Notably absent in the 1940s were systems to coerce drinkers into treatment, which became popular during the unbridled expansion of for-profit treatment facilities. The motivational emphasis in the 1940s, as in the early decades of AA, was on voluntary acceptance of one's condition and of help (Tiebout, 1949).

The treatment systems of the 1940s were also being criticized for ignoring individuals with less severe alcohol problems. Bacon (1949) called for the recognition and treatment of "early stage" incipient alcoholics and non-addicted chronic problem drinkers. He argued that treatment of such individuals could be highly effective. Several writers were calling for differential diagnosis of types of alcoholics, and appropriately different treatments for these subtypes (Bacon, 1949; Bowman & Jellinek, 1941). "Assuming our inability to properly select which patient will be cured by any

particular therapy," Voegtlin and Lemere (1942) asserted, "it would be advantageous to expose the alcoholic first to the well-tested methods of therapy that involve the expenditure of the least time, effort and expense and which are most readily adaptible to treatment of such patients in wholesale fashion" (p. 797). Very similar arguments persisted — equally unheeded — through the 1980s. A large body of research finding no effectiveness advantage for residential over less costly outpatient approaches (Annis, 1986; Miller & Hester, 1986b; U.S. Congress Office of Technology Assessment, 1983) has been met with the perseverating assertion that hospitals are still the best place for alcoholics (Mazza & Stevens, 1988), an unsubstantiated claim that was rejected four decades ago by Jellinek and the World Health Organization.

TREATMENT MODALITIES

In examining the literature of the 1940s, I was impressed: **1.** that a large array of different treatment modalities were available; **2.** that most of them had received little or no scientific evaluation; and **3.** that things again haven't changed as much as one might wish. The principal debate of the decade seems to have been whether treatment should be focused on medical (e.g., medications and hospitals) or nonmedical (e.g., psychotherapeutic and outpatient) strategies (Bowman & Jellinek, 1941).

Almost all of the major treatment modalities of the 1980s were already in use in the 1940s. Kersten (1949) described an inpatient alcoholism treatment program at the Winter V.A. Hospital as consisting of AA meetings on the ward as a "mainstay," twice-weekly group therapy sessions, individual counseling, and the creation of a therapeutic milieu. The only difference from our (Miller & Hester, 1986a) listing of the typical elements of U.S. treatment in the 1980s is the absence of the now common educational lecture series, and the prescription of Antabuse — an element that was added to the Winter V.A. program in the 1950s (Wallerstein *et al.*, 1957).

Group therapy seems to have been a particularly popular, albeit unevaluated approach (Pfeffer, Friedland, & Wortis, 1949). The *Quarterly Journal of Studies on Alcohol* devoted an unprecedented 242 pages to the publication — beginning in 1946 (McCarthy, 1946) and extending over five years — of the verbatim transcripts of 11 group psychotherapy sessions. Individual counseling and psychotherapy approaches of the day apparently emphasized either emotional appeals for abstinence, or the readjustment of a presumed underlying pathological personality (Bowman & Jellinek, 1941; Voegtlin & Lemere, 1942). Though the terminology has changed, these two themes persist (e.g., confrontation, co-dependence) in contemporary alcoholism counseling. Then as now, client motivation

was commonly mentioned as a problem and obstacle in therapy (e.g., Glueck, 1942; Durfee, 1946).

Other contemporary approaches have clear parallels in the literature of the 1940s. Bowman and Jellinek describe "substitutive" strategies (which would now be called broad spectrum therapies) including involvement of clients in alternative hobbies, recreation, occupations, or religious activities. Durfee (1946) described a motivational induction procedure not unlike "motivational interviewing" (Miller, 1983), to be used in the early phase of treatment. He advised that "The therapist must listen with patience, with understanding, and without interruption" (p. 232). Bowman and Jellinek (1941) called for the use of family therapy: "Not only must interference on the part of relatives be eliminated, but they must be integrated into the therapeutic scheme" (p. 158).

Some modern treatments, however, were nowhere apparent in the 1940s. Behavioral self-control training with a goal of moderation, for example, did not make its appearance until 1970 (Lovibond & Caddy, 1970). Some, like covert sensitization (Elkins, 1980), represent newer forms of older strategies — in this case, aversive counterconditioning (Voegtlin & Broz, 1949). Other strategies, such as apneic aversion and the use of metronidazole or LSD in treating alcoholics, arose after the 1940s and already have justifiably disappeared into the pages of history (Miller & Hester, 1980).

I believe that progress *has* been made in specifying methods for carrying out particular treatment modalities. Over the past few decades, treatments called by the same name have often been administered with great variation in procedures. The careful specification and standardization of treatment protocols is an important advance toward understanding treatment efficacy and the processes underlying it. Elkins (1980), for example, described in detail a theory-based set of procedures he used to conduct covert sensitization. This approach had been applied in previous studies, the reports of which usually contained too brief a description to permit replication. Elkins designed his treatment procedures to yield respondent conditioning and avoidance learning, presumed mechanisms for efficacy of aversion therapies (Rimmele, Miller & Dougher, 1989). Applying his procedures, we were able to replicate many of Elkins' findings quite closely, including the number of trials to conditioning and the predictive relationship of conditioning to treatment outcome (Miller & Dougher, 1989). When covert sensitization has been applied in a manner unlikely to produce conditioning (e.g., by group administration), the procedure has been unsuccessful (Sanchez-Craig & Walker, 1982; Telch, Hannon, & Telch, 1984). The clear specification of procedures thus seems to be an important matter. Treatment procedure outlines and manuals are increasingly available as guides for therapists in applying

specific procedures (e.g., Sanchez-Craig, 1984). Through research, steps have also been taken toward specifying more effective approaches in therapist style (Miller, 1983; Miller & Sovereign, 1989) and client/treatment matching (McLellan, Woody, Luborsky, O'Brien, & Druley, 1983).

A variety of medical treatment strategies were in use in the 1940s. Aversion therapy, especially chemical aversion to induce nausea, was offered by the Shadel Hospital (and still is by the same hospital, now owned by Schick). Psychotropic medications (including morphine, barbiturates, and amphetamines) were employed to provide patients with an alternative source of feelings of well-being — a strategy similar to contemporary uses of methadone with heroin addicts or anti-anxiety medications with alcohol abusers. Disulfiram (Antabuse) was introduced as a new treatment agent in the late 1940s (Glud, 1949).

To be sure, there were treatments in use in the 1940s that would, from current perspectives, be judged as bizarre. These included spinal tap, the elevation of blood sugar, insulin shock therapy, prefrontal lobotomy, the injection of serum from intoxicated horses, and the administration of a plethora of drugs including strychnine and gold salts (Voegtlin & Lemere, 1942). (I might mention that on a 1989 trip to the Soviet Union, our delegation encountered a similar array of hair-raising treatments for alcoholism, including cryogenic freezing of the brain, induction of high fever and sweating, dialysis and radiation of the blood, hyperbaric chamber confinement, and electrical stimulation of the brain.) One wonders which of our own currently accepted 'treatments' of the 1980s will be candidates for the chamber of horrors in the 21st century.

TREATMENT OUTCOME REACH

How could such treatments be used? The alcoholism treatment systems of the 1940s (like those of the 1980s) permitted the use of treatments substantiated by nothing more than the belief of the provider in their efficacy. Voegtlin & Lemere (1942), summarizing their review of alcoholism treatment, noted: "The most striking observation is the apparent reticence with which the English-speaking psychiatrists have presented statistical data concerning the efficacy of treatment" (p. 795).

The absence of outcome research is often accompanied by authoritarian assertions of the superiority of favored approaches. Alcoholics Anonymous, for example, was asserted to have "astounding results" (Heath, 1945), an unqualified claim that continues to be echoed four decades later. Unevaluated treatments for alcoholism were promulgated and practiced. Although treatment outcome research was rare, biomedical studies flourished, focusing on the physiological effects of alcohol as a drug, and searching for a biological cause of alcoholism

(e.g., Smith, 1949).

What has changed in 40 years? The search for the biomedical magic bullet continues apace, along with scrutiny of alcohol's pharmacologic effects on virtually every cell system of the body. What is new, over the past three decades, is a relative explosion of treatment assessment research. More than 200 controlled clinical trials have evaluated the efficacy of at least 35 different treatment modalities (Holder, Longabaugh, Miller, & Rubonis, 1991). Over 400 other reports have provided uncontrolled outcome data for a still larger array of alternative interventions (Miller & Hester, 1980). Our professional reluctance to report data has all but vanished.

On the down side, however, it is difficult at present to discern any significant impact of these 40 years of research on current practice (Miller, 1987). The modal alcoholism treatment program continues to rely on modalities that are either inadequately evaluated or unsupported by available research (Holder *et al.*, 1991; Miller & Hester, 1986a). Unsubstantiated pronouncements and practices still abound, but now in the presence (and often defiance) of abundant outcome data. Bowman and Jellinek estimated the average abstinence rate, two to four years after alcoholism treatment, to be 25-30 per cent in 1941. It is discouraging that four decades later, well-designed prospective outcome studies reflect a similar rate of abstinent outcomes among unselected alcoholics (Helzer *et al.*, 1985; Polich, Armor, & Braiker, 1981; Vaillant, 1983). Private programs treating highly selected populations with good prognosis, of course, report somewhat higher abstinence rates (e.g., Wallace, McNeill, Gilfillan, MacLean, & Fanella, 1988), almost identical to those for similar populations of the 1940s (e.g., Voegtlin & Broz, 1949). But after 40 years of research, it remains true that the most appropriate parting words at discharge are not "good bye" but "auf Wiedersehen." A plausible explanation of this inertia is the alcoholism treatment field's intransigence and unresponsiveness to the corrective feedback available through outcome research.

THROUGH A CRYSTAL BALL DARKLY

But you did not invite me here to cast gloom and reflect only on our lack of achievement. What grounds are there for optimism? I believe there are several.

First, I believe that we *already* have at our disposal a very promising armamentarium of treatment alternatives that have been shown, in properly controlled trials, to be effective in ameliorating alcohol problems (Hester & Miller, 1989). To be sure, we have no panacea, no approach that will cure most of the people most of the time. Rather we have a menu of promising options, each of which may be optimal for a definable subset of those with alcohol problems. Neither are we

helpless in choosing from the menu. The 1940s offered a hurly-burly of treatment options, too, but without empirical maps to guide their use. It is apparent now that all treatments are *not* equally effective. Among controlled comparisons of alternative modalities, 43 per cent have yielded positive findings, a significant difference in efficacy between approaches (Holder *et al.*, 1991). Even on the basis of overall track records, there are grounds for making intelligent choices among alternative treatments.

What is still more encouraging is the evolution of methodologies and algorithms for matching individuals to optimal treatments. This is not a new idea. Bowman and Jellinek were calling for individualization of treatment before the Addiction Research Foundation was born. What *is* new is the present array of clever approaches for managing such matching: the cafeteria plan (Ewing, 1977), the core/shell model (Glaser *et al.*, 1984), blocking designs (McLellan *et al.*, 1983), multivariate modeling and self-matching (Miller & Hester, 1986c). There are already encouraging data to support the idea that appropriate matching can dramatically improve treatment outcome (e.g., McLachlan, 1974). Major new studies are underway to strengthen the knowledge base from which to tailor treatment to the individual client. Happily, perhaps not coincidentally, this development rides the crest of what Toffler (1980) has called the "third wave" in Western society — the burgeoning of individualization. Standardization appears to be on the way out, and personal choice seems to rule the day in the media, fashions, recreation, the world of work, and health (cf. Deci & Ryan, 1985; Toffler, 1980). We seem to be headed toward a society in which treatment is designed to fit the individual.

As we are moving toward treatments that are both more effective and more individualized, it also seems likely that the normative treatments of the future will be less intensive, expensive and intrusive. The urgent economics of health care cost containment necessitate a move away from the 1980s' love affair with residential treatment. No overall advantages have been found for residential over nonresidential settings (Miller & Hester, 1986b), although the cost differential is substantial. If for no other reason, economics are likely to require a much more selective use of cost-intensive forms of care in the future.

Meanwhile, a variety of clinical trials have demonstrated beneficial impact from very brief interventions, especially with less severe problem drinkers. Alcohol is joining diet, smoking, and exercise as part of society's rising health consciousness. The realm of secondary prevention has already spawned a WHO collaborative study (Babor, Korner, Wilber, & Possidente-Good, 1987; Saunders, 1987), treatment approaches (Alden, 1988; Brown, 1980; Sanchez-Craig, Annis, Bornet, & MacDonald, 1984), and community-targeted interventions (Heather, 1989) designed to reduce the

consumption of at-risk drinkers. With this development, health care for problem drinkers is escaping from the confines of intensive treatment for severely dependent alcoholics.

Will it happen? Will my optimistic prophecy be fulfilled, that we will move toward a more individualized, cost-effective and broader array of treatment options that serve the full spectrum of people with alcohol problems at lower cost? Will the toil of treatment outcome researchers begin to have a significant impact on practice?

Or perhaps, at your 80th anniversary in 2029, will someone else be standing here, quoting this 40-year-old paper, and reflecting on how very little has changed in the treatment of alcohol problems?

Personally, I remain optimistic. After all, in the words of Warren Wager (1963), "The ultimate function of prophecy is not to tell the future, but to make it."

REFERENCES

Alden, L. E. (1988). Behavioral self-management controlled-drinking strategies in a context of secondary prevention. *Journal of Consulting and Clinical Psychology, 56*, 280-286.

Annis, H. M. (1986). Is inpatient rehabilitation of the alcoholic cost effective? Con position. *Advances in Alcohol and Substance Abuse, 5*, 175-190.

Babor, T. F., Korner, P., Wilber, C., & Possidente-Good, S.(1987). Screening and early intervention strategies for harmful drinkers: Initial lessons from the Amethyst Project. *Australian Drug and Alcohol Review, 6*, 325-339.

Bacon, S. D. (1949). The administration of alcoholism rehabilitation programs. *Quarterly Journal of Studies on Alcohol, 10*, 1-47.

Blumestone, E. M. (1944). Institutional facilities for the treatment of alcoholism: Foreword. *Quarterly Journal of Studies on Alcohol, 5*, 5-8.

Bowman, K. M., & Jellinek, E. M. (1941). Alcohol addiction and its treatment. *Quarterly Journal of Studies on Alcohol, 2*, 98-176.

Brown, R. A. (1980). Conventional education and controlled drinking education courses with convicted drunken drivers. *Behavior Therapy, 11*, 632-642.

Corwin, E. H. L., & Cunningham, E. V. (1944). Institutional facilities for the treatment of alcoholism. *Quarterly Journal of Studies on Alcohol, 5*, 9-85.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Durfee, C. H. (1946). Some practical observations on the treatment of problem drinkers. *Quarterly Journal of Studies on Alcohol, 7*, 228-239.

Elkins, R. L. (1980). Covert sensitization and alcoholism: Contributions of successful conditioning to abstinence maintenance. *Addictive Behaviors, 5*, 67-89.

Emrick, C. D. (1974). A review of psychologically oriented treatment of alcoholism. I. The use and interrelationships of outcome criteria and drinking behavior following treatment. *Quarterly Journal of Studies on Alcohol, 35*, 523-549.

Ewing, J. A. (1977). Matching therapy and patients: The cafeteria plan. *British Journal of Addiction, 72*, 13-18.

Finney. J. W., Moos, R. H., & Mewborn, C. R. (1980). Posttreatment experiences and treatment outcome of alcoholic patients six months and two years after hospitalization. *Journal of Consulting and Clinical Psychology, 48*, 17-29.

Glaser, F. B., Annis, H. M., Skinner, H. A., Pearlman, S.,Segal, R. L., Sisson, B., Ogborne, A. C., Bohnen, E., Gaxda, P., & Zimmerman, T. (1984). *A system of health care delivery.* (3 vols.) Toronto: Addiction Research Foundation.

Glud, E. (1949). The treatment of alcoholic patients in Denmark with "Antabuse": with suggestions for its trial in the United States. *Quarterly Journal of Studies on Alcohol, 10*, 185-197.

Glueck, B. (1942). A critique of present day methods of treatment of alcoholism. *Quarterly Journal of Studies on Alcohol, 3*, 79-91.

Heath, R. G. (1945). Group psychotherapy of alcohol addiction. *Quarterly Journal of Studies on Alcohol, 5*, 555-562.

Heather, N. (1989). Brief intervention strategies. In R.K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches*: Effective alternatives (pp. 93-116). New York: Pergamon.

Heather, N., & Robertson, I. (1983). *Controlled drinking*. (Rev.ed.). London: Methuen.

Heather, N., Whitton, B., & Robertson, I. (1986). Evaluation of a self-help manual for media-recruited problem drinkers: Six month follow-up results. *British Journal of Clinical Psychology, 25*, 19-34.

Helzer, J. E., Robins, L. N., Taylor, J. R., Carey, K., Miller, R. H., Combs-Orme, T., & Farmer, A. (1985). The extent of long-term moderate drinking among alcoholics discharged from medical and psychiatric treatment facilities. *New England Journal of Medicine, 312*, 1678-1682.

Henderson, Y. (1940). The high proof of liquor as a factor in the production of alcoholism. *Quarterly Journal of Studies on Alcohol, 1*, 1-12.

Hester, R. K., & Miller, W. R. (Eds.) (1989). *Handbook of alcoholism treatment approaches: Effective alternatives*. New York: Pergamon.

Holder, H. D., Longabaugh, R., Miller, W. R., & Rubonis, A.V. (1991). The cost effectiveness of treatment for alcohol problems: A first approximation. *Journal of Studies on Alcohol, 52*, 517-540.

Huss, M. (1849). *Alcoholismus chronicus. Chronisk alkoholisjukdom: Ett bidrag till dyskrasiarnas kanndon*. Stockholm: Bonnier/Norstedt.

Jellinek, E. M. (1946). Phases in the drinking history of alcoholics. *Quarterly Journal of Studies on Alcohol, 7*, 1-88.

Jellinek, E. M. (1960). *The disease concept of alcoholism*. Highland Park, NJ: Hillhouse Press.

Kersten, P. M. (1949). Changing concepts in alcoholism and its management. Quarterly Journal of Studies on Alcoholism, *9*, 523-531.

Lovibond, S. H., & Caddy, G. (1970). Discriminated aversive control in the moderation of alcoholics' drinking behavior. *Behavior Therapy, 1*, 437-444.

Manson, M. P. (1948). A psychometric differentiation of alcoholics from nonalcoholics. *Quarterly Journal of Studies on Alcohol, 9*, 175-206.

Mazza, D., & Stevens, P. (1988). Comment on Miller and Hester's "Inpatient alcoholism treatment: Who benefits?" *American Psychologist, 43*, 199-200.

McCarthy, R. G. (1946). Group therapy in an outpatient clinic for the treatment of alcoholism. *Quarterly Journal of Studies on Alcohol, 7*, 98-109.

McLachlan, J. F. C. (1974). Therapy strategies, personality orientation and recovery from alcoholism. *Canadian Psychiatric Association Journal, 19*, 25-30.

McLellan, A. T., Woody, G. E., Luborsky, L., O'Brien, C.P., & Druley, K. A. (1983). Increased effectiveness of substance abuse treatment: A prospective study of patient-treatment 'matching.' *Journal of Nervous and Mental Disease, 171*, 597-605.

Miller, W. R. (1976). Alcoholism scales and objective assessment methods: A review. *Psychological Bulletin, 83*, 649-674.

Miller, W. R. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy, 11*, 147-172.

Miller, W. R. (1987). Behavioural alcohol treatment research advances: Barriers to utilization. *Advances in Behaviour Research and Therapy, 9*, 145-164.

Miller, W. R., & Dougher, M. J. (1989). Covert sensitization: Alternative treatment procedures for alcoholism. *Behavioural Psychotherapy, 17*, 203-220.

Miller, W. R., & Hester, R. K. (1980). Treating the problem drinker: Modern approaches. In W. R. Miller (Ed.), *The addictive behaviors: Treatment of alcoholism, drug abuse, smoking, and obesity* (pp. 11-121). New York: Pergamon.

Miller, W. R., & Hester, R. K. (1986a). The effectiveness of alcoholism treatment: What research reveals. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors: Processes of change* (pp. 121-174).

Miller, W. R., & Hester, R. K. (1986b). Inpatient alcoholism treatment: Who benefits? *American Psychologist, 41*, 794-805.

Miller, W. R., & Hester, R. K. (1986c). Matching problem drinkers with optimal treatments. In W. R. Miller & N. Heather (Eds.), *Treating addictive behaviors: Processes of change* (pp. 175-203).

Miller, W. R., & Sovereign, R. G. (1989). Effects of therapist style on resistance and outcome among problem drinkers. Manuscript in preparation, University of New Mexico.

Pfeffer, A. Z., Friedland, P., & Wortis, S. B. (1949). Group psychotherapy with alcoholics. *Quarterly Journal of Studies on Alcohol, 10*, 198-216.

Polich, J. M., Armor, D. J., & Braiker, H. B. (1981). *The course of alcoholism: Four years after treatment.* New York: Wiley.

Rees, J. R. (1932). Psychological factors in the prevention and treatment of alcohol and drug addiction. *Lancet, 223*, 929-933.

Rice, O. R. (1943). Religion and the church in relation to alcohol addiction. *Quarterly Journal of Studies on Alcohol, 3*, 393-399.

Rimmele, C. T., Miller, W. R., & Dougher, M. J. (1989). Aversion therapies. In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives*, pp. 128-140. New York: Pergamon.

Rush, B. (1943). An inquiry into the effects of ardent spirits upon the human body and mind with an account of the means of preventing and of the remedies for curing them. Reprinted in the *Quarterly Journal of Studies on Alcohol, 4*, 321-341.

Sanchez-Craig, M. (1984). *Therapist's manual for secondary prevention of alcohol problems: Procedures for teaching moderate drinking and abstinence.* Toronto: Addiction Research Foundation.

Sanchez-Craig, M., Annis, H. M., Bornet, A. R., & MacDonald, K. R. (1984). Random assignment to abstinence and controlled drinking: Evaluation of a cognitive-behavioural program for problem drinkers. *Journal of Consulting and Clinical Psychology, 52*, 390-403.

Sanchez-Craig, M., & Walker, K. (1982). Teaching coping skills to chronic alcoholics in a coeducational halfway house: I. Assessment of programme effects. *British Journal of Addiction, 77*, 335-50.

Saunders, J. B. (1987). The WHO project on early detection and treatment of harmful alcohol consumption. *Australian Drug and Alcohol Review, 6*, 325-339.

Smith, J. J. (1949). A medical approach to problem drinking. *Quarterly Journal of Studies on Alcohol, 10*, 251-257.

Straus, R. (1946). Alcohol and the homeless man. *Quarterly Journal of Studies on Alcohol, 7*, 360-404.

Telch, M. J., Hannon, R., & Telch, C. F. (1984). A comparison of cessation strategies for the outpatient alcoholic. *Addictive Behaviors, 9*, 103-110.

Tiebout, H. M. (1949). The act of surrender in the therapeutic process. *Quarterly Journal of Studies on Alcohol, 10*, 48-58.

Toffler, A. (1980). The third wave. New York: Morrow.U. S. Congress Office of Technology Assessment (1983). *The effectiveness and costs of alcoholism treatment.* Washington, D.C.: U.S. Government Printing Office.

Vaillant, G. R. (1983). *The natural history of alcoholism: Causes, patterns, and paths to recovery.* Cambridge, Mass.: Harvard University Press.

Voegtlin, W. L., & Broz, W. R. (1949). The conditioned reflex treatment of chronic alcoholism. X. An analysis of 3125 admissions over a period of ten and a half years. *Annals of Internal Medicine, 30*, 580-597.

Voegtlin, W. L., & Lemere, F. (1942). The treatment of alcohol addiction: A review of the literature. *Quarterly Journal of Studies on Alcohol, 2*, 717-803.

Wager, W. W. (1963). *The city of man.* Baltimore, MD: Penguin Books.

Wallace, J., McNeill, D., Gilfillan, D., MacLean, K., & Fanella, F. (1988). Six-month treatment outcomes in socially stable alcoholics: Abstinence rates. *Journal of Substance Abuse Treatment, 5*, 247-252.

Wallerstein, R. S., Chotlos, J. W., Friend, M. B.,Hammersley, D. W., Perlswig, E. A., & Winship, G. M. (1957). *Hospital treatment of alcoholism: A comparative experimental study.* New York: Basic Books.

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

The Law in Addictions: Principles, Practicalities and Prospects

PATRICIA G. ERICKSON

ADDICTION RESEARCH FOUNDATION

TORONTO, ONTARIO

PATRICIA G. ERICKSON

Dr. Erickson joined the Addiction Research Foundation (ARF) as a scientist in 1973, following several years' research experience with the Ontario Institute for Studies in Education and the Centre of Criminology, University of Toronto. Her undergraduate and Master's degrees in sociology are from York University and the University of Toronto, and she received her doctorate in Criminology/Social Administration from the University of Glasgow, Scotland in 1983. She has written and spoken widely on drug policy and is now a senior scientist in the ARF's Prevention and Health Promotion Research and Development department. Her books include *Cannabis Criminals*, *The Steel Drug* and *Illicit Drugs in Canada*. Recently, Dr. Erickson has been conducting a longitudinal study of cocaine and crack users in the community and developing a public health perspective on illicit drug control.

> "For the criminal law, at least, man has an inalienable right to go to hell in his own fashion, provided he does not directly injure the person or property of another on the way. The criminal law is an inefficient instrument for imposing the good life on others."

(Morris and Hawkins, 1970:2; *The Honest Politician's Guide to Crime Control*)

> "The criminal law may properly be applied, as a matter of principle, to restrict the availability of harmful substances, to prevent a person causing harm to himself or to others by the use of such substances, and to prevent the harm caused to society by such use.... Generally speaking, the adverse effects for the individual of the criminal law process are such that it must be justified in each case by rational and convincing reasons of necessity, in relation to other available means of achieving the desired purpose."

(Le Dain Commission, 1973:940; *"Whether, in principle, the criminal law should be used in the field of non-medical drug use.")*

INTRODUCTION

To inquire about the relationship between drugs and the law, one logical question is: What does knowing a drug's legal status tell you about its pharmacological properties or its dangers? If a drug is *prohibited*, it may be a stimulant, depressant, analgesic or hallucinogen (e.g., cocaine, cannabis, heroin, LSD). If a drug is *legally available*, it may also fall into a variety of pharmacological categories (e.g., nicotine, alcohol, codeine). To reverse the question: what does knowing a drug's effects on the health of a population tell you about its legal status? Those drugs whose known long-term effects are harmful to a significant proportion of users are likely to be legal: "The drugs carrying the greatest health and safety risks are not, in fact, illegal." (*ARF Annual Report*, 1987-88:10). If a drug's short-term effects are potentially lethal, it may be medically prescribed (e.g., barbiturates) or illicit (e.g., cocaine, heroin). Since all the examples thus far cited are addictive drugs, in the sense that they can be used compulsively and destructively, the property of addictive potential is no guide to legal status. Thus, no apparently straightforward relationship between a drug's pharmacological properties and effects, and its position inside or outside the ambit of criminal law, exists.

Nevertheless, the field of addictions is divided into licit and illicit drugs. Although we tend to take this as a given,

it is a phenomenon that originated only in this century. At different times since 1900, all the drugs that present society with problems today (plus a few newcomers like LSD, amphetamines and steroids) have been variously, and at different times, freely available, medically prescribed, considered a candidate for prohibition, considered a candidate for legalization, banned, re-legalized, aggressively marketed and severely restricted. If there is one feature that epitomizes the inter-relationship between the law and drugs, it is ongoing controversy as to how they should be controlled.

OVERVIEW

Although looking back over 40 years of research is the focus of this lecture series, this time period is both too long and too short for the topic of drugs and the law. The first criminal statutes defining illicit drug use were spawned at the beginning of the 1900s. Prohibition of alcohol had ended for the last Canadian holdout, Prince Edward Island (P.E.I.), in the year 1948. When the Addiction Research Foundation was created in 1949, federal legal controls for narcotics were in place, the problem of illicit drug use was small, and heroin was the major illicit drug of concern. Alcohol, though not yet tobacco, was seen as a serious health threat. The expansion of the literature on illicit drug use and its control began about 20 years ago, in the wake of the rapid rise of 'new' expressions of illicit drug use (i.e., cannabis, LSD, cocaine) in the 1960s and 1970s.

Three pivotal phases of scientific examination of drugs and the law can be discerned in the past 40 years. Looking back at the early literature (post-1949), articles from the 1950s and 1960s were authored, for the most part, by lawyers, physicians and police officers, later joined by sociologists and psychologists. These books and papers comment mainly on new legal statutes or are opinion pieces arguing the pros and cons of existing laws. The medically supervised distribution of heroin was a major topic. Very little empirical evidence was brought to bear on the issues, and often the writers concluded with a call for more and better education on the topic of illicit drugs.

In the second phase, a period of more intensive scientific consideration was initiated by the Le Dain Commission in Canada, and similar commissions in other countries. In the 1970s, the development of the empirical testing of deterrent effects of criminal sanctions began in earnest. Marijuana and other drug use were among the most frequently investigated 'crimes' in these studies. These investigations are important because they marked a turning point in the willingness to subject the effectiveness of criminal drug law to scientific scrutiny. Such an empirical approach to drug law made it possible to begin a cost-benefit assessment of the impact of the system of legal controls on

drug use and distribution.

Much of this work on deterrence in the 1970s was in the domain of sociologists and criminologists. In the third phase in the 1980s, the literature on the control of addictions has expanded greatly with contributions from economists, political scientists and public health specialists. Cost-benefit models became comprehensive, incorporating the economic costs of efforts to suppress drug trafficking, the effects on foreign relations, and other crime-related costs, and the measurement of health costs generated by drug use became more sophisticated.

In the discussion to follow, I will address three main points. First, the functions of the law and alternative means of social control will be described. Second, the types of drug law will be reviewed. Third, the effectiveness of the law in controlling drug use will be considered with respect to the lessons provided by alcohol, heroin, cannabis and cocaine.

FUNCTIONS OF LAW

Social control is an essential element of human society: "(it) is found wherever and whenever people hold each other to standards... it divides people into those who are respectable and those who are not; it disgraces some, but protects the reputations of others" (Black, 1976:105). There is no group so pure, so homogeneous, that some individual's infraction of standards cannot be isolated and punished to reinforce the group's cohesiveness and solidarity — so argued Emile Durkheim (1895:100) nearly 100 years ago in his "community of saints" example. Thus, from the Durkheimian perspective, the purpose of the law is to provide a "collective conscience" for society. From a Marxist perspective, laws function to resolve conflicts rooted in class interests by imposing the will of the powerful on subordinate groups. In yet another perspective, the law is seen as a way of balancing competing interests in a pluralistic society. All of these functions of law are illustrated to some extent by the evolution of drug law, a fascinating topic in itself, but one beyond the scope of this paper.[1] (See Musto, 1973.)

Most forms of social control are informal, i.e., enacted through such routes as etiquette, custom, gossip and sermons, and arise spontaneously from individual and group interactions. Social control enacted and enforced by government is law, the most intrusive involvement of the state in the normative lives of its citizens. Law arises, it is thought, when increasing numbers of unrelated people live in close proximity to each other, and traditional forms of conflict resolution become inadequate.

A fundamental principle in the sociology of law is, "Law varies inversely with other social control" (Black, 1989: 74). Put another way, law is only one way to manage conflict and impose control, but the more alternatives there are, the less

law is required, and conversely, the fewer alternatives, the more law. Other options include self-help, avoidance, negotiation, settlement and tolerance (Black, 1989). Most conflicts are handled without law.

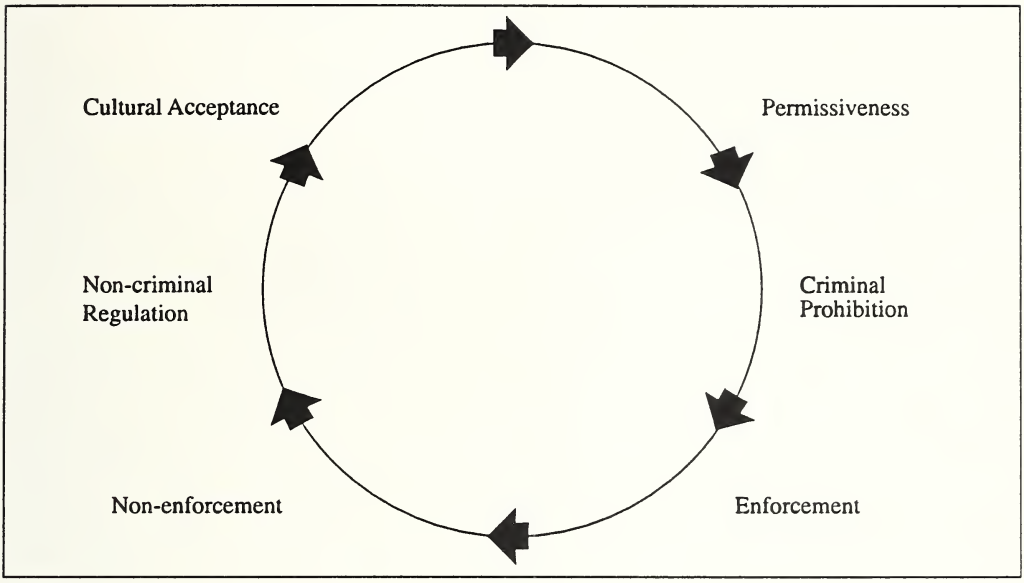
What are some examples of other forms of social control besides law in relation to drug behavior? *Self-help*, in which a grievance is resolved by unilateral aggression, covers a wide spectrum of possibilities. The turf wars over crack in inner cities of Los Angeles and Washington are examples of violent conflict resolution without resort to law. So are many activities of international drug trafficking cartels. Another scenario of self-help is when a youthful drug user is disciplined by an authority figure such as a parent or teacher. *Avoidance*, in which potential interaction is minimized, is illustrated by the persistence of a drug-using community like Rochdale College, which was easily shunned by those who objected to the activities. *Negotiation* is a process by which a compromise is sought, such as when parents and children negotiate certain conditions pertaining to the use of alcohol, tobacco or other drugs inside (or outside) the family home. *Settlement* involves a third party, for instance, the involvement of a drug counsellor or physician in a parent-child dispute over drug use. Settlement sometimes occurs in *absentia*, such as when gossip, which may be an important force in high school groups, ostracizes drug users. The final mode, *tolerance* (or learning to live with it, biting the bullet, letting bygones be bygones, etc.) is a frequent response to the drug use of friends, despite not condoning it.

Law is resorted to less often than other options, even when the acts involved are quite serious and could constitute crimes. A crime is any act legally punishable by the state (Glaser, 1978). But "laws exist merely as a *potential* for sanctioning" (Goode, 1970:284). It is evident from victimization surveys and other self-report measures that about two-thirds of incidents regarded as 'crime' by citizens go unreported to the police (Black, 1989:77). When citizens do not regard acts such as drug use or violence between friends as serious, legal action is even more unlikely.

If law is a relatively unusual form of social control, why has it increased its sway in this century, and why has it been extended to *some* behaviors involving the consumption and distribution of psychoactive drugs? The short answer is that we have come to rely on it because of the perceived absence of alternatives (Black, 1989). As other options became less available in our more fragmented, specialized, heterogeneous culture, we turned more and more to law. As family and communal authority declined, central authority grew. Becoming accustomed to legal redress, we sought other options less.

Over the past 100 years, the criminal law was extended to create new crimes that previously had been controlled less formally. The attitude of many legal critics, in a ground swell of jurisprudential opposition in the 1960s, is reflected in the

Figure 1. Control Cycle: Non-predatory Conduct



titles of their works: *The Limits of the Criminal Sanction* (Packer, 1968), *The Crisis of Overcriminalization* (Kadish, 1967) *The Borderland of Criminal Justice* (Allen, 1964), and *Not the Law's Business* (Geis, 1979). Others in this tradition have suggested, with reference specifically to marijuana, that "the criminal justice system is not the place to demonstrate our sentiments of condemnation" (Hagan, 1979:198). The Le Dain Commission took this approach in considering the appropriate scope of the criminal law and the need to develop alternatives. The point of overdependence on law will be pursued later, after considering the general preventive effect of the criminal law, and then assessing our past and current legal response to drug use. First, how has law evolved, and what are the key concepts in distinguishing types of drug law?

TYPES OF DRUG LAWS

All criminal laws are rooted in conflict and take their shape from the jockeying for influence of various interest groups with diverse objectives (Glaser, 1978). A widespread consensus gradually emerges on the need to punish a variety of *predatory* crimes, in which some harm or loss is imposed on victimized parties. These are what people usually think of as 'crimes': murder, theft, kidnapping, arson, assault, etc. The number of criminal offences deemed predations has expanded enormously since the Middle Ages, and statutes that define victimizing crimes are much more likely to be reformulated than rescinded. Once established, such laws take on a formality and legitimacy in their own right.

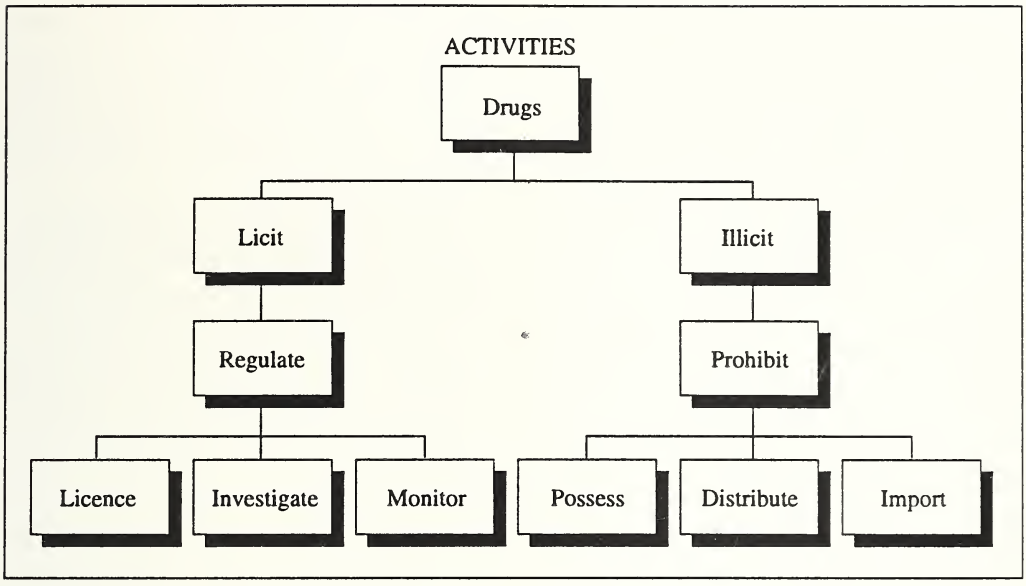
In sharp contrast to the persistence

of predatory crime laws, however, are the ongoing conflicts, suspensions and repeals of numerous laws pertaining to *nonpredatory* behavior. Such offences involve voluntary participation. As such they include the consensual or 'vice' crimes such as prostitution, homosexuality, abortion, drug use, some sexual acts and gambling, some of which are no longer deemed crimes (Cook, 1988). Nonpredatory conduct becomes defined as criminal through the activities of interest groups that first, are threatened by values and lifestyles that conflict with their own and second, are able to mobilize sufficient political power for the passing of laws. Prohibition becomes organized around some standard of acceptable behavior that may be defined quite rigidly; the establishment and persistence of these laws are extended and consolidated by criminal justice bureaucracies (Glaser, 1978:23). Nonpredatory offences are generally difficult to detect and punish, since they are conducted by willing participants, usually in private, and do not result in a complainant (victim). Desirable commodities or services are sought by willing consumers. These crimes are characterized by few arrests in relation to the extent of the activity, selective arrest and prosecution, police reliance on numerous expensive and intrusive tactics (e.g., informants and decoys), and often untoward levels of corruption of officialdom (Kaplan, 1988). State punishment for some nonpredatory offences has declined (e.g., Sunday 'Blue Laws,' gambling, abortion) but has been retained for others such as certain forms of drug use. Legal reaction to much nonpredatory behavior tends to go through cycles, from permissiveness to prohibition to regulation. (See **Figure 1**). Moral ambivalence about these behaviors generates controversy. What then has been the cycle for psychoactive substances?

Drugs in North America were treated much as any other commodity, and their casualties were treated as more deserving of pity and assistance than censure, until early in the twentieth century (Brecher, 1972). Then permissiveness ended, to be replaced by two conceptually distinct legal approaches for the control of drug use, *prohibition* and *regulation*. (See **Figure 2**). At one extreme, prohibition seeks to prevent by suppression all use of, and access to, a drug through criminal laws that create offences (of possession, distribution, manufacture and importation). These crimes are subject to severe penalties for violation. In contrast, regulation allows some acceptable level of use of a drug, by at least some members of society, but access is restricted through various health protection statutes defining the quality of the product, conditions of sale and guidelines for medical prescription.

Violations of regulations are usually punished by fines, or by suspensions or revocations of licenses. In practice, some overlap between prohibition and regulation occurs, as when small amounts of prohibited drugs are available for medical or

Figure 2. Drugs and the Law



research purposes, or when criminal penalties are imposed for the illegal sale, but not possession, of regulated drugs. Use of a licit substance such as alcohol may also be subject to criminal sanction in specific situations, like public drunkenness or drinking and driving (Giffen, 1976). While the agents of proscription — the police, courts and customs officials — are empowered to detect and punish all users and sellers, the agents in a regulatory scheme — the physicians, scientists and bureaucrats — are engaged in investigating, licensing and monitoring the drug's use, purity and availability.

Canada's history in drug control is not atypical of other Western nations. The legal framework of our current system of drug control in Canada was laid down in the early part of this century. By 1908, proprietary and patent medicines, as well as tobacco and alcohol, were on the "road to regulation" (Murray, 1988). The opiates and cocaine were embedded in criminal legislation, the *Opium and Drug Act* of 1911, to be joined by cannabis in 1923. Despite a debate in Parliament over a private member's bill to prohibit tobacco, efforts to criminalize tobacco failed, but temperance groups gained ground during the First World War when all provinces enacted alcohol prohibition (Smart and Ogborne, 1986). By 1929, all provinces except P.E.I. had rescinded their alcohol prohibition statutes and imposed regulation on the trade. International prohibition and regulation through the *Single Convention on Narcotic Drugs and the Convention on Psychotropic Substances* have reinforced this division between licit and illicit drugs in many countries (Smart *et al.*, 1981; 1983).

◆

The types of legal control in effect against a particular drug seems to occur in stages. The sequence of permissiveness, prohibition and regulation can be affected by a variety of factors including the extent of use, the perception of dangerousness, the existence of social movements opposed to particular drug use and who the users are (Glaser, 1974; Skolnick, 1988). Permissiveness towards the use of any drug generally prevails when its effects are not perceived as serious, especially if the drug is taken in a form that does not suggest a medication, e.g., cigarette smoking. If the drug is viewed as a minimally dangerous medication, some minor pharmaceutical restrictions may be imposed (e.g., over-the-counter preparations). Social movements to enact prohibition laws have been successful when a drug is regarded as dangerous (though not necessarily only to health) and morally suspect. Legislation is more readily enacted when the substance is used only by a small, low-status part of the population; the majority are unaffected by and indifferent to the prohibition (e.g., opium in 1908; opiates, cocaine and marijuana from 1911-1929). If larger, more influential segments of the population oppose prohibition, either the movement was unsuccessful (e.g., tobacco) or legislation did not survive long (e.g., alcohol in the post-World War I period). The state then may move to regulation, whereby the licensing of availability minimizes some of the social costs of unrestricted access and use. Alcohol laws in Ontario, for instance, were quite restrictive after prohibition, but then became progressively more liberalized (Single *et al.*, 1981). Tobacco, on the other hand, has moved of late from permissiveness to greater regulation in Canada and some other countries.

Opposition to the prohibition of cannabis, the opiates and cocaine has shown different patterns. When a drug becomes much more widely used, as happened to cannabis in the 1960s and 1970s, enforcement is highly selective (as it was for alcohol during prohibition) and this perceived injustice fuels repeal arguments, particularly if substantial numbers of the middle class or their children are being criminalized. The recognition that marijuana was not as harmful as it had been portrayed earlier in the century also contributed to the opposition to total prohibition (Himmelstein, 1986). The next stage for cannabis, however, halted short of regulation, at penalty reduction. Increased emphasis on cannabis' potential adverse effects and its perceived status as a 'gateway' drug to opiates and cocaine (Clayton, 1985) appeared to stymie any further move for repeal of prohibition or even for penalty reduction in Canada and the majority of American states.

Sometimes the perception of the problem as one of health rather than morality increases prior to any retreat from criminalization. Thus, opiate addicts became subject to civil commitment, and (while prohibition was retained) methadone programs became available in the U.S. Some European countries took

a more regulatory approach to supplying narcotics on a restricted and controlled basis to opiate addicts. When drugs are widely used in legitimate medicine, but also dangerous and subject to misuse, strict medical regulation is likely to be the primary method of control, with criminal penalties available for unauthorized manufacture and distribution but not possession (e.g., amphetamines, barbiturates, tranquillizers). Coexistence or overlap between stages in the cycle of control is possible, and no one stage may prevail permanently. What lessons have been learned from the experience of legal control through prohibition of various drugs over several decades? How has the criminal law been applied, with what intent and what effect?

DETERRENCE: A SYSTEM OF GENERAL PREVENTION

The usual North American term, 'general deterrence,' or its European counterpart 'general prevention,' refers to both the perceived and demonstrated ability of the criminal law and its enforcement to make citizens law-abiding (Andenaes, 1975). The deterrent effectiveness of our legal system has long been (for practitioners at least) an article of faith that is central to the philosophy of criminal law, legislation and sentencing:

> "Judges, politicians and even lawyers... view [research on deterrence] as a needless demonstration of the painfully obvious."
> (Gibbs, 1973, quoted in Andenaes, 1975:338)

Similar assumptions have been accepted by those considering general prevention as applied to illicit drugs:

> "It is impossible to determine with any certainty or precision how far a criminal law is having deterrent effect. Our continuing resort to such laws... is based on the assumption that they have deterrent effect... in the absence of convincing proof or arguments to the contrary. *We cannot put into question the assumption which underlies the whole of our criminal law.*"
> (Le Dain, 1972:289; emphasis added)

Yet the philosopy of general deterrence makes implicit assumptions and hypotheses about human behavior, namely that lawbreaking is inversely related to the certainty, severity and swiftness of punishment (Zimring and Hawkins, 1973). Not until the 1970s were there extensive empirical examinations of general deterrence. Underlying this effort was the recognition that since the presumed success of deterrence can only be measured in terms of non-events, i.e., illegal acts that did *not* occur, its effectiveness is

Figure 3. Deterrence

- GENERAL DETERRENCE
  - Threat to potential violators
  - Perceived: CERTAINTY, SEVERITY, SWIFTNESS
  - Moral-Educative guidelines for community
- SPECIFIC DETERRENCE
  - Prevents further offence by the individual
  - Demonstrated: CERTAINTY, SEVERITY, SWIFTNESS
  - Makes the threat credible to the community
  - Reaffirms guidelines for the community

difficult to gauge. Moreover, since officially recorded crimes do not necessarily correlate with the population's knowledge of the certainty and severity of legal penalties, the emphasis shifted from studies of fluctuation in crime rates to the perceptions of legal threats by actual and would-be offenders.

The mechanisms assumed to underly general prevention can take several forms. (See **Figure 3**). The threat of punitive consequences is meant to discourage people from law-breaking. Those who commit crimes are to be punished (i.e., specific deterrence), both as a lesson to refrain from future infractions, and to make the threat credible to the populace at large. The educative mechanism sets standards for undesirable behavior, placing boundaries between what will be merely disapproved of and what may be subject to criminal sanction. A related component is the symbolic, moral disapproval of certain acts. Thus the concept of general prevention extends beyond the threat of legal punishments to include the law's ability to strengthen inhibitions against criminal acts by performing an educative and moralizing function.

When the law is invoked as a means of social control, it is important to distinguish intent from effect. This leads to several major questions: To what extent is people's conduct directed by the threat of punishment? What are the prerequisites for an effective legal prohibition? Can a prohibition be maintained that is not reflective of the public's moral code? What happens to the moral authority of law when prohibitions are openly disregarded? The answers to these questions require objective assessment of the available evidence. If these questions can be answered, a rational discourse on deterrent effectiveness can set the stage for addressing dispassionately the separate ethical and political issues of when to invoke the general preventive option of criminal law.

LIMITS OF THE LAW

Introduction

With the emphasis on current drugs of concern, the prohibition on alcohol is a distant memory, and efforts to suppress caffeine and nicotine are an even more remote part of Western history. It is easy to overlook the fact that all forms of drug use were confined to, and part of, indigenous cultures before international trade, and especially before the great European trading expansion of the 15th and 16th centuries. Europe was rapidly converted from a culture that used alcohol only, to one that used many drugs (Brecher, 1972).

Coffee, for example, had met with considerable resistance and suppression when first introduced into Arabia and Egypt. Despite the punishment of users and the destruction of supplies, it persevered to become the national beverage of much of the Arabic world. Although coffee became a popular drink in Europe too, bars that served coffee and hot chocolate were looked upon with suspicion, perhaps because of the political dissidents they were perceived to attract. Medical warnings of the dangers of coffee addiction also proliferated. Cases of coffee psychosis were cited, coffee drinkers were viewed as prone to turn to "other narcotics," and as late as 1909, a medical textbook warned about "coffee excess" as a cause of extreme mental and physical misery (Brecher, 1972:198). Currently, despite sporadic reports about the adverse effects of excessive caffeine intake, this drug, in the form of tea, coffee, cocoa and cola drinks, has become the drug of choice for a broad span of age groups, and control of its use is a matter of personal or parental regulation.

Tobacco, smoked in pipes and cigars or snorted as snuff, spread rapidly as both crop and commodity from the New World to Western societies during the 16th and 17th centuries. Until supply caught up with demand, tobacco was expensive, being matched with silver, ounce for ounce, in the London of the early 1600s, where 7,000 tobacco shops flourished (Brecher, 1972:210). In this era, a number of countries and municipalities issued edicts prohibiting smoking, in locales as far removed from each other as Bavaria, Russia and Japan. The death penalty was decreed for tobacco use by the Sultan of Constantinople in 1633. Yet in no country where tobacco use was established did it ever lose ground. The modern cigarette was not marketed in quantity until the end of the 19th century, and quickly became the most popular way to take a 'hit' of nicotine. Possibly because tobacco became more accessible in this form to women and children, and partly in keeping with the temperance sentiments of the era, several concerted efforts were made to prohibit cigarette smoking. In the 1920s, some states of the U.S. enacted short-lived cigarette prohibitions, but these soon were repealed due to the economic clout of producers. What remained were laws against sale to minors, much as are in

effect in the U.S. and Canada today. The most recent efforts to restrict smoking have taken the form of regulatory laws directed at advertising and also at smoking in public places, vehicles and working environments. Smoking remains a personal choice, albeit an increasingly constrained one, as the private dwelling and the great outdoors become almost the last refuges of the nicotine addict.

These historical examples of failed prohibitions against caffeine and nicotine tell us much about the difficulty in suppressing the widespread use of drugs that are popular in all strata of society, but tell us little about how prohibition might operate within a modern criminal justice context. For those lessons, it is necessary to turn to alcohol, heroin, cannabis and cocaine, as each became the target of intensive suppression in different decades of the 20th century.

1. LESSONS FROM ALCOHOL: 1920s AND 1930s

The first large scale experiment with prohibition in the modern world was with alcohol.[2] National prohibition was in effect in the U.S. from 1919 to 1933, and in Canada alcohol was banned from 1915 to 1918 in all provinces. Since anti-alcohol laws pertained mainly to production and sale, and consumers were not the targets of enforcement, this era might more properly be described as 'partial prohibition,' akin to the decriminalization of marijuana possession in more recent times (Single, 1989). Or, since alcohol was available on prescription in Ontario, with physicians issuing nearly one million prescriptions in 1923-24, 'prohibition' was also similar to the controlled drug status of amphetamines and barbiturates in the contemporary *Food and Drugs Act* (Smart and Ogborne, 1986). Nevertheless, alcohol prohibition has become the prototype of the failure of criminal law in the addictions field. Several classic studies have been done in the U.S. (e.g., Gusfield, 1963) and new work has been published on the Canadian experience with prohibition (Smart and Ogborne, 1986). Although prohibition is now assessed in terms of its public health impacts, the impetus was the moral pressure of temperance ideology, not a concern for health (Schweighofer, 1988; Smart and Obgorne, 1986).

For most close observers of that era, prohibition did not likely completely 'fail' in the sense of a limited objective of reducing overall alcohol consumption, but was rather a costly mistake. Research evidence from this era is limited and largely impressionistic. Indirect evidence from Ontario showed a 40 per cent drop in alcohol-related deaths (including cirrhosis) during prohibition, compared to the previous five years (Smart and Ogborne, 1986:59). It is also possible, however, that reporting practices changed, as would be expected when behaviors became illegal. Adverse health effects were experienced by people who drank impure or

contaminated alcohol from local producers. What became quickly evident was that the laws were not only largely unenforceable and widely broken, but also generated high costs in terms of corruption, violence and the fostering of large-scale criminal organizations.

Drinking did continue, in somewhat different outlets and settings. While taverns and bars were closed, illegal drinking places (speakeasies) flourished 24 hours a day. In Canada, each province could manufacture alcohol 'for export' and much of this supply was transferred between provinces or went to the U.S. One of the first results of reduced availability of a desired commodity is an increase in price. The bootlegging fraternity was expanded by those eager to share in the enormous profits, and those arrested were readily replaced by new manpower. Some of the profits were channeled to local police and customs officials who would conveniently look the other way as required. For example, boats bringing liquor from Windsor, Ontario to nearby Detroit were registered by officials as having gone to Cuba. One small motorboat operating out of Windsor was shown as having exported a boatload of liquor from there to Cuba four times in one day (Smart and Ogborne, 1986:54). Reports on the corruption and dishonesty of the enforcers helped to turn public opinion against the laws banning alcohol.

The level of violence escalated, as rival gangs competed with each other for the profitable trade; bootleggers also had confrontations with police and special agents. The necessity of 'stool pigeons' and 'spotters' to spy on the public, and the use of open search warrants, further alienated the majority of otherwise law-abiding citizens. Violence in the illicit trade seemed to be more extreme in the U.S. than in Canada; Smart and Ogborne (1986) note that carrying guns and 'shooting it out' was much less common north of the border. In the U.S., the level of violence and corruption of officials is epitomized by the Al Capone era of Chicago. The very administration of justice in that city became compromised by a sort of 'mob rule.' Here too the infamous St. Valentine's Day Massacre illustrated the extent to which large-scale powerful criminal organizations would go in securing their monopoly. Innocent victims and enforcement officials added to the toll of violence. Although the Chicago experience came to epitomize the costs of prohibition, Kaplan (1988) has commented that the very integrity of U.S. government institutions generally was compromised by the unprecedented levels of corruption.

Within a relatively few years, national and local prohibition lost ground. What had been voted in was voted out. The moral ambivalence of the Prohibition Era had been nicely captured by Will Rogers: "If you think this country ain't Dry, you just watch 'em vote; and if you think this country ain't Wet, you just watch 'em drink" (quoted in Cook, 1988:11). Prohibition was repealed, and was replaced by a system of regulation.

Repeal occurred not because alcohol was seen as a safe or harmless drug, but because harsh lessons were learned about the limitations of the law in preventing the use of an already desired and well-established commodity. Calls for more resources and powers devoted to the enforcement of prohibition were shown to be ineffective and contributed to the increasing lawlessness of the trade. Powerful economic interest groups saw the advantage of legally producing alcohol, high status members of the society wished to consume alcohol without inconvenience and governments saw the benefits of taxing alcohol. In the end, Al Capone was brought down by tax law — not anti-alcohol law. But what if the substance of concern is restricted to a much smaller and powerless group?

2. LESSONS FROM HEROIN: 1950s AND 1960s

The heroin addict had been the prototype 'junkie' or 'dope fiend' since the advent of prohibition, when the addict-criminal replaced the addict-patient (King, 1974). As a number of pioneer researchers documented in classic works, the major lesson of heroin prohibition was that it generated a particularly damaging lifestyle for addicts (Stephenson, 1956; Chein *et al.*, 1964; Lindesmith, 1965). Many of the unhealthy and socially destructive conditions displayed by addicts popularly attributed to the opiates' actions on the mind and body were shown to be, instead, the results of the narcotics laws and the social conditions they produce.

When the U.S. *Harrison Act* was passed in 1914, there were about one-quarter million opiate addicts in the U.S. About two-thirds were women; use was associated with being middle-aged, white and from the middle class; most led normal, productive lives (King, 1974). Opiates were cheap and readily available, and addicts did not, nor were perceived to, pose a threat to the community. While estimating precisely the number of practitioners of illegal habits is notoriously difficult, the National Institute of Mental Health [NIMH] placed the figure for 1971 at one-quarter million opiate addicts in the whole country (Brecher, 1972:62). This may have been an underestimate, as according to King (1974), there were an estimated 100,000 addicts in New York city alone, and possibly close to one-half million in the whole country in the early 1970s.[3] Most addicts by then were young, poor, and members of ethnic or social minorities. The addict lifestyle had become associated with predatory crime, and though the victims were far more likely to be members of the addicts' ghetto communities, the perceived threat pervaded all segments of American society. What had happened in the interim? I will limit discussion to only a few key points, but the research in this area is extensive (see Inciardi and Chambers, 1974). It is also of great current

importance in understanding the transmission of the AIDS virus among intravenous drug users.

In 1962, a U.S. Supreme Court decision characterized opiate addicts as "the walking dead," and cited a list of horrors associated with heroin users, including emaciation, dental disease, jaundice, tuberculosis, bronchitis, boils and abscesses, collapsed and scarred veins and insanity (Brecher, 1972:21). This stereotype was important in maintaining the status quo of prohibition because, as Duster (1970:29) notes, "many of the intensely moralistic judgements about narcotics are based on conceptions of the physiological consequences of the drugs." But investigators of that era could find "not even one scientific study on the proven harmful effects of [opiate] addiction" (Stephenson, 1956:509). Instead, a number of studies and sources from past years provided quite a contrasting view:

> "Morphine addiction is not characterized by physical deterioration or impairment of physical fitness aside from the addiction per se." (1929)
>
> "The addict when not deprived of his opium showed no abnormal behavior which distinguished him from a non-addict." (1931)
>
> "The addict under his normal tolerance of morphine is medically a well man." (1932)
>
> "Given an addict who is receiving [adequate] morphine... the deviations from normal physiological behavior are minor." (1940)
>
> "The data of this study indicate that the habitual use of morphine does not cause a chronic psychosis or an organic type of deterioration." (1946)
>
> [all cited in Brecher, 1972:21-27]

More recent medical and psychiatric commentators continued to agree that the overall effects of opiates on the minds and bodies of addicts were not directly harmful, and were indeed much less damaging than the chronic use of alcohol and tobacco.

> "Chronic psychoses as a result of the excessive use of opiates are virtually nonexistent." (Kolb, 1962:120)
>
> "The addict who is able to obtain an adequate supply of drugs through legitimate channels usually... remains in good health, suffers little inconvenience, and is, in general, difficult to distinguish from other persons." (Jaffe, 1970:286)

> "Opiate addiction in and of itself is not physically dangerous... Ingesting opiate extracts by mouth is safer than injecting morphine, but if dosage is reasonably well-controlled, even opiates by injection are not inherently dangerous." (Snyder, 1989:38)

If it was not the drugs themselves, then what could account for the poor health and degraded lives of the heroin addicts?

Many heroin researchers of the 1950s and 1960s probably started on their studies with the assumption that the law was a necessary deterrent to this dangerous practice. Close examination of the problem, however, led to a different perspective. Here, the work of Chein *et al.* (1964) and Lindesmith (1965) is reflective of many other investigators who reached similar conclusions.

> "We did not start [the research project] with any suspicion that the law might be contributing to the problem... we did not, then, have any notion that the most dangerous consequences of addiction to the individual were a direct outcome of the existence and the enforcement of the law." (Chein *et al.*, 1964:6-7)

To summarize a vast amount of consistent research findings: forced to obtain opiates illegally and at a much inflated price, addicts formed a deviant and sometimes criminal subculture; money went to purchase drugs rather than to food and personal hygiene; sharing and using dirty needles led to high rates of infection with diseases such as hepatitis; contaminated drugs and unknown potency contributed to high morbidity and mortality; criminalization further reduced legitimate economic opportunities. In terms of the social threat and burden posed by addicts, the relatively inelastic demand for heroin, coupled with the high price of maintaining a habit, led a substantial proportion of addicts to engage in predatory street crime (this relationship has received much detailed analysis in more recent years [see Anglin and Speckart, 1988]).

As the numbers of heroin addicts increased in the postwar period, the social response was to toughen the legal penalties and increase enforcement. As Chein *et al.* (1964:8) noted, "the inflation of the addiction problem to a national menace... then justifies harsher penalties, more restrictive measures and more restrictions on individual rights." Both the U.S. and Canada toughened their narcotics laws in the 1950s and 1960s: mandatory minimum sentences were introduced, maximum sentences were increased to life imprisonment, narcotics law violators were denied eligibility for parole, laws were redefined to place the onus of proof on the defendant rather than the prosecution, and trafficking was broadly defined to include giving and sharing — common practices amongst

opiate users. Police manpower assigned to narcotics work was also increased. Rarely was enforcement successful against the upper echelons of the drug distribution trade, only a small fraction of the traffic was intercepted (estimates even then were five to 10 per cent; see Brecher 1972:93), and the black market continued to flourish. Successful drug interceptions would reduce supplies, raise risks, raise prices, increase profits for the retailer and generate more crime by addicts.

Those who bore the brunt of the criminal law were the street addicts (not the medical or society ones[4]) who were also usually small-time dealers. Stephenson's account (1956:415) shows an appreciation of the minimal deterrent effects of the law on addicts and the hardships wrought. He and his colleagues estimated, from a variety of sources, that 900 active heroin addicts were living in British Columbia in 1955 (excluding health professionals). Over a 20-month period, this group received a total of 755 prison sentences. Some spent the whole period in prison and others were sentenced two or three times. On an average day, 550 of the province's street addicts were in prison, leaving 350 at large to occupy the 32 full time narcotics officers — one officer for every 11 addicts. Yet the number of addicts did not decline, and most continued a cycle of arrests after release. Even if personality "weaknesses were shown to pre-date addiction, years of crime, years of prison, years of unemployment, years of anti-social hostility (and society's anti-addict hostility).... these can hardly be expected to strengthen a person-ality and eradicate its weaknesses" (Stephenson, 1956:518).

This was not an era of one-sided debate on narcotics policy. Some resistance was voiced to the trend of increasing severity in the name of deterrence. In New Jersey, Governor Robert Meyer used his veto against bills in 1956 that would have increased penalties for narcotics violations, arguing that the tenet of these proposals is that "Injustice to the individual is a fair price to pay for the deterrent effect upon the addict as well as the non-addict sellers. We do not sacrifice individuals for the common good; *rather we find the common good in doing justice to the individual.*" (Lindesmith, 1965:29, emphasis added). The British approach of allowing medical practioners to prescribe heroin for addicts was scrutinized and methadone maintenance was introduced in the U.S. and Canada. Nevertheless, arguments for more humane treatment of addicts, for adequate treatment facilities, for greater consideration of individual rights, fell for the most part on deaf ears. The highly negative stereotype of the addict, and the acceptance of criminal controls, continued to dominate public discourse. The accumulation of scientific evidence appeared to have little impact when the basis for sanctions was strongly-held moral disapproval. The fact that the addicts were a remote and powerless group (small in Canada, ghettoized in

the U.S.) provided little impetus for a major rethinking of society's response to opiate use.

Since opiate use is an infrequent occurrence among the general population, and hence difficult to accurately record in surveys, less is known about the prevalence of initiation, continuation and cessation of heroin use than about those who become addicted (Blackwell, 1983). Studies of the preventive effect of the law have investigated mainly the supply side (Wagstaff and Maynard, 1988; Reuter and Kleiman, 1986), or documented the failure of specific deterrence among opiate addicts and the need for alternative strategies to reduce the incidence of HIV infection. But what if the drug in question is more widely used and thus its devotees are more accessible to the researcher?

3. LESSONS FROM CANNABIS: 1970s

The new wave of cannabis use that began in the 1960s and continued an upward swing through the 1970s provided the opportunity to test deterrence hypotheses in representative population samples where (at least for those under 25 years) nearly equal proportions had, and had not, engaged in this activity. The accompanying rise in criminalization for cannabis offences (mainly possession) generated research on the social effects of punishment on drug users who, unlike the heroin addict subgroup, did not form a deviant, low-status segment of the population. The findings of various studies from the 1970s combined to show a low deterrent benefit coupled with high social costs (Solomon, Single and Erickson, 1983).

Empirical research on general deterrence flourished in the 1970s. Using self-reports from large samples of young people, numerous researchers assessed the relative impact of perceived certainty and severity of punishment (i.e., formal controls) and extra-legal influences such as social background, social support, moral commitment and perceived health risks (i.e., informal controls) on compliance with the law. The thrust of all these findings was to demonstrate the very weak role of legal threats compared to extra-legal factors in decisions to use cannabis (see Erickson, 1980:86-91, for a review). Marijuana use became the most studied crime in the deterrence literature, enabling one investigator who exploited it for a comparative analysis of methodologies to conclude that however and wherever studied, "perceptions of formal sanctions play little or no role in explaining variance in rates of self-reported marijuana use" (Lundman, 1986:381).

These findings on perceptions of legal risks were paralleled by survey research in the U.S. that indicated that in the 1970s, rates of increase of marijuana use did not vary between states that had reduced penalties and those that had not (Johnston, 1980; Single, 1989). Coupled with the low *objective* likelihood of punishment, it is clear that perceptions

were not unrealistic (about one in 100 of those individuals who *used* in the past year were arrested); the proportion of use *incidents* that are punished relative to all those committed would be even smaller, likely one in several thousand (Erickson, 1980:88). Silberman (1976) also reported that while moral commitment could be an important intervening variable between perceived certainty of arrest and law-breaking, the legal prohibition against marijuana was denied legitimacy; use was subject more to social support than to condemnation, thus further eroding deterrent effects. A study of the deterrent impact of the law on marijuana sellers showed a marginal rather than an absolute effect — that is, they sold less frequently because of fear of arrest, but did not stop (Ekland-Olson *et al.*, 1984).

If deterrence was ineffective in stemming increased marijuana use in the 1970s, was it perhaps partly responsible for the decline in the 1980s registered in both the U.S. and Canada? This seems unlikely for several reasons. First, objective properties of certainty and severity did not alter during this decade. Second, perceived risk of getting caught did not increase during the period when use decreased (Smart and Adlaf, 1989). Indeed, heavier users were found to have a greater perception of arrest risk than lighter users (Erickson, 1989). Third, both greater social disapproval of marijuana use and greater perceived risk of harm from marijuana use were found to account for a substantial proportion of the decline (Bachman, Johnston, O'Malley and Humphrey, 1988). Thus, it seems reasonable to conclude that legal threats were as remote a force in the trend of declining use as they were in the upswing of the previous decade and a half. Studies of long-term cannabis users — those who persisted in the practice through the 1970s and 1980s — indicate that their disregard for the law was unchanged (Erickson, 1989). What is not clear, however, is the extent to which the criminal law has exerted a less direct, moral and educative role in reinforcing the greater unacceptability of cannabis among would-be users.

A second lesson from cannabis in the 1970s was the considerable costs that prohibition generated in terms of the criminalization of hundreds of thousands of mainly youthful users in Canada (Le Dain, 1972), and millions in the U.S. (Shafer, 1972). An extensive empirical, longitudinal study of such a group documented that adverse individual consequences did occur, and with a broad range of impact (Erickson, 1980). While most of this sample of 95 first offenders for cannabis possession resisted criminal self-identification, they feared that others such as parents and employers might stigmatize them. The unknown future consequences of a criminal record placed them in a vulnerable position. The selective process of drug enforcement ensured that those officially labelled as "cannabis criminals" were among the youthful, more disadvantaged segment of the cannabis user population (Meier and

Johnston, 1977). Thus, for most offenders, the record was an additional handicap to future prospects. The disruptions caused by arrest and court appearance tended to be transitory and idiosyncratic, involving mainly stress with family members. Some disrespect for police and courts was generated by the firsthand experience of harsh or seemingly arbitrary treatment. The absolute discharge did not mitigate most of the harmful effects of criminalization, and exacerbated the perceptions of uneven sentencing. On the whole, these adverse consequences were not compensated for by specific deterrent benefits: 92 per cent of the offenders continued to use one year after court, and those who stopped were much more likely to be the lighter or experimental users at the time of arrest. A replication study seven years later showed no major changes in the characteristics of cannabis offenders or the impact of criminalization on them, but the police and courts had become more efficient in processing offenders and reducing the burden on the criminal justice system (Erickson and Murray, 1986; Murray and Erickson, 1983). Aging cannabis users learned to "live with prohibition" (Erickson, 1989).

In the 1980s, cannabis use, while persisting at fairly high levels, was replaced in public discourse by the re-emergence of a 'new' drug of concern.

4. LESSONS FROM COCAINE: 1980s

To the lessons already learned from the prohibition of alcohol, heroin and cannabis, was added evidence of the immense power and intractability of the international cocaine supply network. The violence and corruption already endemic in the illicit drug market were amplified, both in the consuming and the producing countries (mainly in the American inner cities and in Colombia). International relations were affected and whole governments destabilized (Bagley, 1988). Despite the recency of the newest wave of cocaine use and legal response, a number of scholars have addressed the issues from a variety of disciplinary perspectives (Wisotosky, 1986; Erickson, Adlaf, Murray and Smart, 1987; Reuter, 1988; Bagley, 1988). New laws related to the seizure of property of suspected drug criminals and to the detection of money laundering of drug profits were passed in a number of countries, and international cooperation in these areas was heightened. What have been the lessons to date?

The first and most important observation about cocaine in the 1980s has been the growth of the illicit cocaine trade. Whether concealed in nativity scenes, hollowed-out navel oranges, or any other of the hundreds of possibilities utilized by smugglers, the amount of cocaine hydrochloride shipped to the U.S., as well as to Canada and newer markets, has increased. Efforts to stop shipments at the borders have met with very limited success, with most enforcers' estimates converging around five to 10 per cent of imported drugs being seized (Wisotsky, 1986).

Nevertheless, expenditures of more resources for interdiction in the past decade (from $1 billion to nearly $4 billion in the U.S.) have been ineffective in preventing greater availability of cocaine on the street. The higher purity of samples seized, and declining prices, have been among the most notable and widely documented results of cocaine prohibition. Although prices vary from city to city, the reported median price for a kilogram of cocaine at the end of 1985 was $34,000 in Florida, $38,000 in New York and $40,000 in the rest of the nation (Reuter, 1988:237). More recently, in New York, the retail price of a kilo of cocaine dropped from $60,000 in 1980 to $12,000 in 1988 (*The Economist*, 2 Sept., 1989:21). Of course, the profits even in the time of the cocaine glut remained huge: broken into four doses of crack per gram, the $12,000 investment could fetch $40,000 on the street.

The reasons for the failure of interdiction to keep prices high, the traditional justification for spending so much effort on this activity, has been explored thoroughly by Reuter (1988). Briefly, the production costs of cocaine at source (the South American countries growing and refining it) are very low compared to the retail price. Thus, traffickers can replace the cocaine seized at the border readily at about one per cent of the retail cost. That means that the largest seizure to date, that of 20 tons of cocaine worth an estimated $2 billion dollars in California in 1989, can be replaced by $20 million dollars of investment at source. The traffickers can easily pass on additional costs, such as when pilots or captains demand extra pay to compensate for increased risks, to the consumer. Reuter (1988) also speculates that more players who are able to avoid detection have entered the trafficking scene, and interviews with convicted high-level dealers suggest that skill requirements are meagre in comparison to the substantial profits to be made (Reuter and Haaga, 1989).

Certainly the greater availability of cocaine within a country such as the U.S. has increased the competition in the local market. This has led to more inner-city violence, as drug dealers vie for control of turf. The dramatic increases in urban murder rates are usually due to dealers and gang members killing each other, but also include the deaths of innocent bystanders, referred to as "mushrooms" by the gangs for their tendency to pop up unexpectedly. Police officers too, always vulnerable in undercover situations, have been added to the tallies of drug-related killings. As noted in the earlier examples of drug prohibition, violence is endemic in the illicit market both because of the kinds of criminals that are attracted to it and because there are no legal institutions to resolve disputes (Goldstein, 1989). All that is 'new' about cocaine in this regard has been the intensity, particularly in the United States and in Colombia.

Another by-product of prohibition, the corruption of police, has also received a great impetus from the enormous profits

of the cocaine trade. Dombrink (1988), in describing a number of incidents of drug-related police corruption, has referred to them as the "Touchables." In Florida, one of the major entry points for large shipments of cocaine, the temptations are considerable and omnipresent. The reality is that the price of a kilogram of cocaine in the 1980s was about the equivalent of a Miami police officer's annual salary (Dombrink, 1988). In this city, a ring of police officers was tried for running a cocaine distribution enterprise. Not only local police have been involved; an FBI agent pled guilty in Miami to accepting cocaine kickbacks; in fact, a special FBI corruption squad has been operating for several years there, directed in large part at drug-related police corruption. Similar cases, of police extorting money from crack dealers, and of police officers stealing, selling and using cocaine, abound wherever the drug market is concentrated. Again, it is not that activities are new, but that the range and opportunities have been expanded by the burgeoning cocaine trade.

Even more serious and graphic accounts of violence and corruption have been provided in the country of Colombia. A number of foreign affairs experts (Bagley, 1988; Lee, 1988; Sharpe, 1988; Hartlyn, 1988; Tokatlian, 1988) have analyzed the impact of the cocaine trade, and efforts to suppress it, on the governments and economies of the major Latin American countries involved. The entrenchment of the cocaine cartels, and their campaigns of terror against elected government officials, the press, and judicial figures, have outstripped any previous images of the extensive costs inherent in prohibition. The destabilization of governments, the impairment of American foreign relations in South and Central America, and the stresses on international relations, have become recognized as indirect consequences of the immense wealth and power generated by the illicit trade in cocaine. An awareness of the ineffectiveness of supply-side enforcement and interdiction has fueled the more recent emphasis on demand-side measures in the consuming countries. Of course, these can take the form of heightened criminalization of users, as illustrated in the recent American Bush-Bennett proposals (White House, 1989), or other less coercive means emphasizing health promotion and education as in Britain and Holland (Bagley, 1988).

The recognition that the cocaine drug barons and their higher chain of command represented the "new untouchables," virtually immune from traditional methods of drug enforcement and prosecution, has prompted the development of strategies for financial controls. If the supply of the drug could not, for the most part, be seriously impeded in its journey from the jungles of Peru and Bolivia and the Colombian processing laboratories, perhaps the supply of money back to the upper echelons of the trade could be interrupted. These attacks on finances took two main forms, namely, the seizure and forfeiture of assets and the tracing of money flow ('money laundering'). It is early yet to assess their effectiveness,

but some preliminary evidence suggests a fairly limited impact. For instance, in Canada the RCMP's anti-drug profiteering program reports an increase every year in assets seized, from $8.4 million in 1983 to $18.6 million in 1987, for a five-year total of $51 million (RCMP, 1987/88). This nonetheless represents a fairly tiny proportion of the RCMP's estimate of an illicit drug trade worth $10 billion in 1984 alone ($1.2 billion for cocaine [RCMP, 1984/85]). Other independent estimates of the value of the Canadian illicit drug market have been considerably lower, in the $2 billion to $3.2 billion range. But even the lower estimate would indicate that less than one per cent of drug profits are seized. A similar estimate has been arrived at for the U.S. program (Wisotsky, 1986:80). The other major financial initiative, the tracing of 'dirty money,' has been a slow and cumbersome procedure, but has produced results in prosecutions of banks. Fines in the millions of dollars against offending banks may be expected to have some deterrent impact on legitimate banking institutions (e.g., $2.75 million against the Crocker National Bank of San Francisco for failure to report $3.9 billion in cash transactions; Wisotsky, 1986:82). Wisotsky (1986) suggests, however, that as long as cash can be smuggled to safe tax havens in other countries, money laundering will be difficult to impede and will be readily absorbed by traffickers as one more relatively minor risk of doing business.

The notion of 'money laundering' took on a somewhat different meaning with the report from the chief toxicologist in Miami who found measurable traces of cocaine on 131 of 135 small denominational bills taken from banks in 12 cities. The average amount per bill was a seven-millionth of a gram, meaning that 5,000 bills would produce one line of cocaine. If all the 12 billion bank notes circulating in America bore these seven micrograms, then the total amount of cocaine actually circulating in the currency would amount to 84 kilos (*The Economist*, April 15, 1989:32).

Insofar as the law might be considered to have any deterrent effect on the cocaine user, studies have indicated that this is minimal. As with cannabis, health threats are more salient (Erickson *et al.*, 1987). A comparison of perceptions of cannabis and cocaine in an adult sample who used both drugs showed that cocaine was seen as more subject to social disapproval, posing greater risks to health and carrying a higher potential for addiction (Erickson and Murray, 1989). A study conducted in the mid-70s had showed that cannabis users had a fairly positive attitude to trying cocaine (Erickson, 1982). It appears that with the passage of time, knowledge of cocaine's risks became more widespread despite any change in law or legal responses. Since cocaine users form a much smaller proportion of the population than cannabis users, broader studies of deterrent effects in the general population are not as useful in studying legal impacts. Since virtually

all cocaine users are previous or current cannabis users, a sense of insulation from legal consequences is likely to typify this group as well.

DISCUSSION AND CONCLUSION

What does the law do well and poorly in its deterrent capacity? What are the prospects for its role in the future?

The criminal law, as the Le Dain Commission and the Law Reform Commission of Canada have emphasized, has costs. These costs are considerable to the individual offender who experiences some loss of control over his/her life and carries the stigma of a criminal record; to society, which pays the social and economic costs of enforcing the prohibition; and to all members of society whose liberty and choices are to some extent constrained by the criminal law. If the criminal law, then, is meant to reduce the harm caused by drug use, it is not unreasonable to require that demonstrable benefits that outweigh the costs should accrue.

The criminal law is a system of general prevention. The Le Dain Commission provided the opportunity to look empirically at the impact of the law. If, as I have suggested, the law is based on certain principles of human behavior, and hypotheses as to behavior change, then can we not adopt the same analytic (and inherently skeptical) approach we apply to other issues in the field of addictions? Simply put, the objectives of using the law in the control of drug use must be clear, realistic, achievable and evaluated. On this basis, it becomes possible to know what to do next that might improve the attainment of objectives.

The guiding principle of the use of the criminal law as a primary instrument of social control of drug use is its general preventive effect. If functioning ideally, the deterrent threat of punishment in combination with the law's educative-moral effects will ensure a high degree of law-abiding behavior and agreement on the necessity of punishing transgressors. The evidence reviewed about the actual practices associated with efforts to suppress drug use and distribution in a prohibitionary scheme illustrates, however, considerable noncompliance with the law, divergent viewpoints on its appropriateness, plus associated high levels of individual and social costs. Why is the law not more effective in shaping behavior in this particular area?

First, deterrent effects are supposed to be raised by certainty and severity of punishment. But the relevant research indicated that drug users perceive the risks of arrest as remote, though greater than those perceived by non-users, and the objective reality of detection of drug activities is, in fact, low, in keeping with their status as nonpredatory crimes. Moreover, contrary to the postulates of deterrence, frequent drug users have a

higher perception of the risk of punishment than do lighter users. On a related point, the fear of stigma following arrest and conviction for possession has been eroded as a deterrent by the increasing 'normalization' of the offence, at least for cannabis users. Severity of punishment likely loses some of its relevance when actual and perceived certainty of detection in the first place is so remote. Legal knowledge among drug users, though somewhat more accurate than in the population-at-large, is inadequate enough to show a wide gulf between intended threats and their perceptions by target groups.

The educative or socializing effects of the law, by setting out standards of acceptable behavior, are meant to encourage compliance. This role is undermined, in practice, by the contrast between legal categories and other knowledge relevant to drug problems. For example, cannabis, cocaine and heroin belong to different pharmacological categories and carry different health risks, yet in Canada all are legally 'narcotics.' Prohibited drugs are no more addictive than their restricted counterparts (amphetamines and barbiturates) or their regulated competitors (alcohol and nicotine). Added to this seeming illogic (as perceived by users and potential users) is the severity of the maximum sentences possible, and which seem out of proportion to any harm caused directly to others by the use of these particular prohibited drugs, as compared to the legal ones. Thus the credibility of the law as a guide to acceptable behavior is beleaguered by such inconsistencies.

The moral authority of the law is activated in part through its presumed impartiality. This is weakened by the widespread violations of the drug laws with little enforcement relevant to the total volume of drug law-breaking that occurs. Differential arrest and prosecution of drug offenders — i.e., many higher status violators go unpunished — further brings the authority of the law into question. When awareness both of the nonenforcement and of the discrepancy in what enforcement occurs becomes widespread, disrespect for the law is increased.

Another aspect of the moral function of the law is the provision of the opportunity for normative validation, or the shared condemnation of certain acts. This stance is more difficult to sustain in a pluralistic society, when opinion is sharply divided on the 'rightness' or 'wrongness' of certain acts. It is difficult to view behavior that is engaged in by a substantial proportion of the population as *seriously* criminal. The laws may be retained as a largely symbolic and ineffectual gesture of disapproval, as long as actual enforcement is minimal. Drug users who do not accept the authority of the law over this aspect of their behavior can readily rationalize their own choice without feeling 'criminal.' Thus the drug laws have failed to command the consensus much more readily extended to predatory crimes against person or property. If the practicalities of prohibition suggest that the law is a

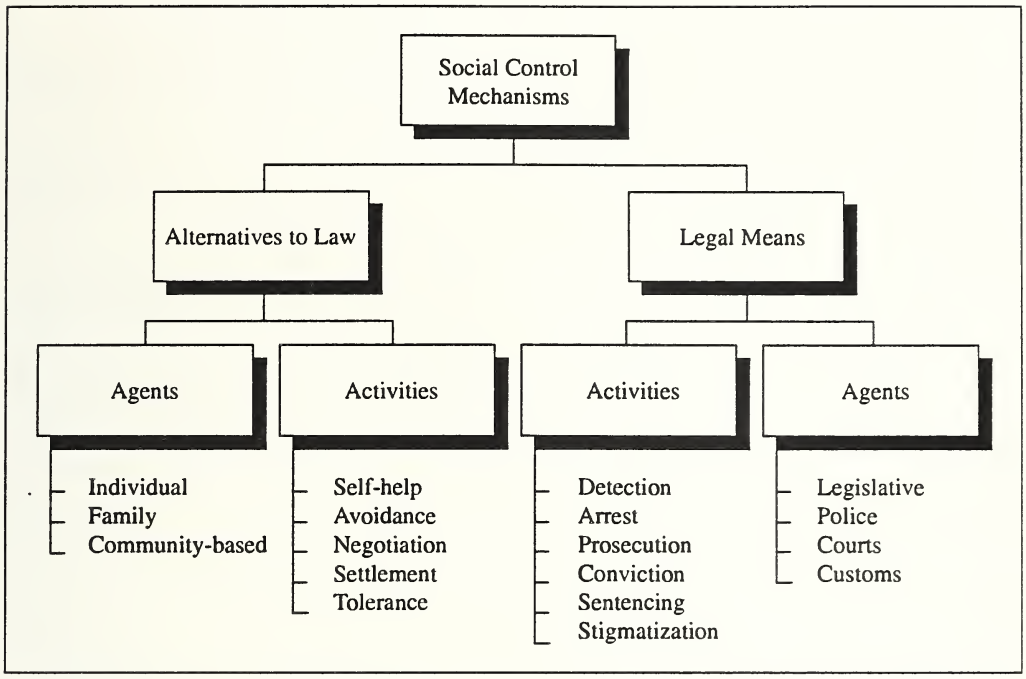
weak form of social control over drug use, why then do we continue to rely on it so extensively, and what are the future prospects for the prevention of drug problems?

If other forms of prevention were completely effective, the law would not be needed. But, if law is inversely related to other forms of social control, then it is likely that gradual relaxation of legal controls may be needed to stimulate the development of alternatives. Conversely, the development of other controls may reduce the amount of legal intervention. There is no disputing that drug use can be dangerous and poses risks that society must endeavor to minimize, but at an acceptable level of social costs. Thus the law may be applied where it is likely to have the most effectiveness (for example, situational specific incidents like use in public or when operating a vehicle) where it can command credibility, and can receive the greatest public support. The enforcement of prohibition may drain resources to activities with little payoff in terms of minimizing harm from drug use (e.g., interdiction at borders; street-level dealing) at the expense of more broadly-based prevention efforts that provide alternatives of treatment and community resources. The deleterious effects of prohibition are evident in their impact on users — the unhealthy and degraded lives of heroin users; the criminalization of youthful cannabis users — without the compensation of significant, demonstrable deterrent effects.

Nevertheless, it is also important to recall that the law performs an important social integrative function when it sets the boundaries of moral consensus on acceptable and unacceptable behavior. Thus, when publicity is generated by the arrests and/or sentencing of drug criminals, the collective conscience is expressing the value that, for example, "we will not tolerate dealing in illegal drugs." Much of the 'drug war' rhetoric can be viewed as an effort to maintain or strengthen this moral consensus. This is essentially a retributivist position, since no rehabilitative or deterrent benefit accrues. On the other hand, to rely exclusively or excessively on legal intervention is to achieve a state of legal overdependence, in which the law becomes like an addictive drug (Black, 1989). In part, people do not always know what the alternatives are, or how to activate them (see **Figure 4**). This misplaced focus on law enforcement can be seen as an expression of frustration and impotence. The call for more enforcement and more severe punishment can serve as a distraction from the underlying social conditions that foster drug abuse, and which cannot be addressed by legal intervention.

I would argue that it is important to continue the process of refocusing the social control of drug use in the individual, the family and the community. The key is to regain at these levels more of the responsibility that has been delegated to an abstraction: the law. The criminal law is an inefficient tool for the control of

Figure 4.



nonpredatory behavior. In attempting to impose a moral consensus on a pluralistic society, attention is diverted from what the major causes of drug problems are. These cannot be addressed simply by reactive, punitive measures. As our society has become more fragmented and compartmentalized, the responsibility for enforcement of social order has been given to specialists, i.e., prosecutors, police, sentencers and jailers: Is it time to be reminded that social order is the responsibility of all citizens? The lessons of legal over-dependence, seen in totalitarian regimes (e.g., China, Iran) are sobering. Under such regimes, almost any behavioral expression can be defined as threatening to the state and hence punishable. Citizens spy on citizens. When the law is fully in command, morality itself loses relevance (e.g., Orwell's *1984*).

I think that over the next 40 years, we will see a reduced role for criminal law and the growth of a system of regulatory and other strategies for prevention of drug problems. I think the process of seeking and testing alternative social controls will be enhanced by cross-cultural research, and will evolve to be more community-based and variable rather than centrally determined. I think cooperative international efforts will continue to be directed at the control of production and supply. Ultimately, a long-range shift to regulation will likely occur for some forms of currently prohibited drug activity and these will probably be more influenced by economic and foreign policy issues

◆

than by either health or social concerns. If the process is handled poorly, without sufficient planning, we in the addictions field will be among those picking up the pieces. If it is handled well, perhaps 2029 will see a more rational, effective and humane approach in place.

I would like to close with a quotation from Daniel Glaser, Professor Emeritus at UCLA, and a close observer of drug prohibition for more than 40 years:

> "Literatures on what governments should do about personal vices of all types — from gambling to commercial sex to recreational drugs — evolve in similar directions.
>
> "They begin with cries of moral outrage and exhortations to proscribe; they end with research reports that contrast what ought to be with what can be, and call for regulation rather than probibition.
>
> "Perspectives on gambling, alcohol and prostitution control [for example] have moved further in such an evolution than perspectives on narcotics policies....
>
> "The trend toward greater tolerance for conduct that does not directly [harm] others is probably inexorable in the long run."
>
> (Glaser, 1985:619)

The lessons of prohibition, which have been costly, suggest a future course in which other kinds of controls must be considered. The choice is not between law and no law, or even between regulation and prohibition; rather it is a matter of determining a more effective balance in approaches to social control.

NOTES

[1] See Cook (1969), Boyd (1984), Comack (1985) and Solomon and Green (1988) for different interpretations of the origins of Canadian narcotic laws.

[2] In the early 1800s, China had attempted to ban opium that was brought into China by the British East India Company. The resulting Opium Wars of 1839-42 forced China to continue receiving opium imports (see Waley, 1958).

[3] Current estimates of intravenous drug users in the U.S. run between 1.1 and 1.3 million, with over 200,000 in New York City. In the state of New York, over 35,000 patients are enrolled in methadone maintenance programs (Selwyn *et al.*, 1989).

[4] In Canada, for example, the Bureau of Dangerous Drugs classified "known addicts" as either "criminal," "professional" (those who were members of the health professions) or "medical" (those under treatment for non-addictive diseases); only the socially marginal "criminal" group was the focus of police attention (Giffen and Lambert, 1988).

REFERENCES

Allen, F. A. (1964). *The Borderland of Criminal Justice*. Chicago: University of Chicago Press.

Andenaes, J. (1975). General prevention revisited: Research and policy implications. *Journal of Criminal Law and Criminology*, 66(3):338-365.

Anglin, Douglas M. and George Speckart (1988). Narcotics use and crime: A multisample, multimethod analysis. *Criminology*, 26(2):197-233.

Bachman, Jerald G., Lloyd D. Johnston, Patrick M. O'Malley and Ronald H. Humphrey (1988). Explaining the recent decline in marijuana use: Differentiating the effects of perceived risks, disapproval, and general lifestyle factors. *Journal of Health and Social Behavior*, 29(March):92-112.

Bagley, Bruce Michael (1988). U.S. foreign policy and the War on Drugs: Analysis of a policy failure. *Journal of Interamerican Studies and World Affairs*, 30(2-3)(Summer-Fall):189-212.

Black, Donald (1976). *The Behavior of Law*. New York: Academic Press.

Black, Donald (1989). *Sociological Justice*. New York: Oxford University Press.

Blackwell, J. (1983). Drifting, controlling and overcoming: Opiate users who avoid becoming chronically dependent. *Journal of Drug Issues*, 13(2):159-162.

Boyd, N. (1984). The origins of Canadian narcotics legislation: The process of criminalization in historical context. *Dalhousie Law Journal*, 8:102-136.

Brecher, E. M. [and the Editors of Consumers Reports] (1972). *Licit and Illicit Drugs*. Mount Vernon, New York: Consumers Union.

Chein, I., D. L. Gerrard, R. S. Lee and E. Rosenfeld (1964). *The Road to H: Narcotics, Delinquency and Social Policy*. New York: Basic Books.

Clayton, R. (1985). Cocaine use in the United States: In a blizzard or just being snowed? pp. 8-34 in N. J. Kozel and E. H. Adams, *Cocaine Use in America: Epidemiologic and Clinical Perspectives*, NIDA Research Monograph No. 61. Rockville, Maryland: NIDA.

Comack, A. E. (1985). The origins of Canadian drug legislation: Labelling versus analysis. In: T. Fleming (ed.), *The New Criminologies in Canada*. Toronto: Oxford University Press.

Cook, S. (1969). Canadian narcotics legislation, 1908-1923: A conflict model interpretation. *Canadian Review of Sociology and Anthropology*, 6(1):36-46.

Cook, P. J. (1988). A introduction to vice. *Law and Contemporary Problems*, 51(1):1-8.

Dombrink, John (1988). The touchables: Vice and corruption in the 1980s. *Law and Contemporary Problems*, 51(1)(Winter):201-232.

Durkheim, E. (1895). *The Rules of Sociological Method*. S. Solovay, J. Mueller and G. Catlin (trans. and eds.). New York: Free Press [1938].

Duster, T. (1970). *The Legislation of Morality*. New York: MacMillan.

The Economist (1989). Drugs: It doesn't have to be like this, 2 Sept.:21. Drug Money, 15 April:32.

Ekland-Olson, S., J. Lieb and L. Furcher (1984). The paradoxical impact of criminal sanctions: Some microstructural findings. *Law and Society Review*, 18(2):159-178.

Erickson, P. G.(1980). *Cannabis Criminals: The Social Effects of Punishment on Drug Users*. Toronto: Addiction Research Foundation Books.

Erickson, P. G. (1982). Illicit drug use, peer attitudes and perceptions of harmful effects among convicted cannabis offenders. *International Journal of the Addictions*, 17(1):141-154.

Erickson, P. G. (1989). Living with prohibition: Regular cannabis users, legal sanctions and informal controls. *International Journal of the Addictions*, 24(3):175-188.

Erickson, P. G., E. M. Adlaf, G. F. Murray and R. G. Smart (1987). *The Steel Drug: Cocaine in Perspective*. Lexington: D. C. Heath and Co.

Erickson, P. G. and G. F. Murray (1986). Cannabis criminals revisited. *British Journal of Addiction*, 81(1):81-85.

Erickson, P. G. and G. F. Murray (1989). The undeterred cocaine user: Intention to quit and its relationship to perceived legal and health threats. *Contemporary Drug Problems* 16(2):141-156.

Geis, G.(1979). *Not the Law's Business*. Rockville, Maryland: National Institute of Mental Health, Centre for Studies in Crime and Delinquency.

Giffen, P. J.(1976). The criminal courts and the control of addictions. In: M. Friedland (ed.), *Courts and Trials, A Multidisciplinary Approach*. Toronto: University of Toronto Press.

Giffen, P. J. and S. Lambert (1988). What happened on the way to law reform? pp. 345-369. In: J. Blackwell and P. Erickson (eds.), *Illicit Drugs in Canada: A Risky Business*. Toronto: Nelson Canada.

Glaser, Daniel (1974). Interlocking dualities in drug use, drug control and crime, pp. 39-56. In: James A. Inciardi and Carl D. Chambers (eds.), *Drugs and the Criminal Justice System*, Vol. II. Beverly Hills: Sage.

Glaser, Daniel (1978). *Crime in Our Changing Society*. New York: Holt, Reinhart and Winston.

Glaser, Daniel (1985). The criminal law's nemesis: Drug control. *American Bar Foundation Research Journal*:619-626.

Goldstein, P. J.(1989). *Pathways to Criminal Violence*, pp. 16-48. In: N. A. Weiner and M. E. Wolfgang (eds.). Newbury Park, CA: Sage.

Goode, E. (1970). *The Marijuana Smokers*. New York: Basic Books.

Gusfield, J.(1963). *Symbolic Crusade: Status Politics and the American Temperance Movement*. Urbana: University of Illinois Press.

Hagan, J. (1977). *The Disreputable Pleasures: Crime and Deviance in Canada.* Toronto: McGraw-Hill Ryerson.

Hartlyn, J. (1988). Commentary on Bagley's 'The new 100 years war? National security and drugs in Latin America.' *Journal of Interamerican Studies and World Affairs*, 30(1):183-186.

Himmelstein, J.(1986). The continuing career of marijuana: Backlash... within limits. *Contemporary Drug Problems*, 13(1):1-21.

Inciardi, James A. and C. Chambers (1974). *Drugs and the Criminal Justice System.* Beverly Hills: Sage.

Jaffe, J. H. (1970). In: L. Goodman and A. Gilman (eds.), p. 286, *The Pharmacological Basis of Therapeutics*, 4th Ed. New York: MacMillan.

Kadish, S. H.(1967). The crisis of overcriminalization. *The Annals of the American Academy of Social and Political Science*, 374 :157.

Kaplan, John (1988). Taking drugs seriously. *The Public Interest*, 92(Summer):32-50.

King, R. (1974). 'The American system': Legal sanctions to repress drug abuse, pp. 17-37. In: J. Inciardi and C . Chambers (eds.), *Drugs and the Criminal Justice System.* Beverly Hills: Sage.

Kolb, L. (1962). *Drug Addiction, A Medical Problem.* Springfield, IL: Charles C. Thomas.

Le Dain Commission [Commission of Inquiry into the Non-medical Use of Drugs] (1972). *Cannabis.* Ottawa: Information Canada.

Le Dain Commission (1973). *Final Report.* Ottawa: Information Canada.

Lee, Rensselaer III (1988). Dimensions of the South American cocaine industry. *Journal of Interamerican Studies and World Affairs*, 30(2-3)(Summer/Fall):87-103.

Lindesmith, Alfred (1965). *The Addict and the Law.* Bloomington: Indiana University Press.

Lundman, R. J. (1986). One-wave perceptual deterrence literature: Some grounds for the renewed examination of cross-sectional methods. *Journal of Research in Crime and Delinquency*, 23(4): 370-388.

Meier, R. F. and W. T. Johnson (1977). Deterrence as social control: The legal and extra-legal production of conformity. *American Sociological Review*, 42(April):292-304.

Morris, N. and G. Hawkins (1970). *The Honest Politician's Guide to Crime Control.* Chicago: University of Chicago Press.

Murray, G. F. (1988). The road to regulation: Patent medicines in Canada in historical perspective, pp. 72-87. In: J. Blackwell and P. G. Erickson (eds.), *Illicit Drugs in Canada: A Risky Business.* Toronto: Nelson Canada.

Murray, G. F. and P. G. Erickson (1983). Regional variation in criminal justice system practices: Cannabis possession in Ontario. *Criminal Law Quarterly*, 26(1):74-96.

Musto, D. (1973). *The American Disease: Origins of Narcotic Control*. New Haven: Yale University Press.

Packer, H. L. (1968). *The Limits of the Criminal Sanction*. Stanford: Stanford University Press.

Reuter, Peter (1988). Quantity illusions and paradoxes of drug interdiction: Federal intervention into vice policy. *Law and Contemporary Problems*, 51(1)(Winter):233-252.

Reuter, P. and J. Haaga (1989). The Organization of High-Level Drug Markets: An Exploratory Study. *A RAND Note*. Santa Monica, CA: The RAND Corporation.

Reuter, P. and M. Kleiman (1986). Risks and prices: An economic analysis of drug enforcement, pp. 289-340. In: M. Tonry and N. Morris (eds.), *Crime and Justice: An Annual Review of Research*, Vol. 7. Chicago: The University of Chicago Press.

Royal Canadian Mounted Police [RCMP] (1988). *National Drug Intelligence Estimate* 1986/87. Ottawa: Minister of Supply and Services.

RCMP (1989). *National Drug Intelligence Estimate 1987/88*. Ottawa: Minister of Supply and Services.

Schweighofer, A. R. F. (1988). The Canadian temperance movement: Contemporary parallels. *Canadian Journal of Law and Society*, 3:175-194.

Selwyn, P., D. Hartel, W. Wasserman and E. Drucker (1989). Impact of the AIDS epidemic on morbidity and mortality among intravenous drug users in a New York City methadone maintenance program. *American Journal of Public Health*, 79(October):1358-1362.

Shafer Commission [National Commission on Marijuana and Drug Abuse] (1972). *Marijuana: A Signal of Misunderstanding*. Washington, D. C.: U. S. Government Printing Office.

Sharpe, Kenneth E. (1988). The drug war: Going after supply, a commentary. *Journal of Interamerican Studies and World Affairs*, 30(2-3)(Summer/Fall):77-85.

Silberman, M. (1976). Toward a theory of criminal deterrence. *American Sociological Review* 41(June):442-461.

Single, E. (1989). The impact of marijuana decriminalization: An update. *Journal of Public Health Policy* (in press).

Single, E., N. Giesbrecht and B. Eakins (1981). The alcohol policy debate in Ontario in the postwar era, pp. 127-158. In: E. Single, P. Morgan and J. de Lint (eds.), *Alcohol, Society and the State*, Vol. 2. Toronto: Addiction Research Foundation.

Skolnick, J. H. (1988). The social transformation of vice. *Law and Contemporary Problems*, 51(1):9-30.

Smart, R. G. and E. Adlaf (1989). Student cannabis use and enforcement activity in Canada: 1977-1987. *Drug and Alcohol Dependence* 24(1):67-74.

Smart, R. G., G. F. Murray and H. D. Archibald (1981). *Psychotropic Substances and Their International Control*. Toronto: Addiction Research Foundation.

Smart, R. G., G. F. Murray and H. D. Archibald (1983). *Report of the International Working Group on the Single Convention on Narcotic Drugs, 196l.* Toronto: Addiction Research Foundation.

Smart, R. G. and A. C. Ogborne (1986). *Northern Spirits: Drinking in Canada Then and Now.* Toronto: Addiction Research Foundation.

Snyder. S. (1989). *Brainstorming: The Science and Politics of Opiate Research.* Cambridge, MA: Harvard University Press.

Solomon, R. and M. Green (1988). The first century: The history of non-medical opiate use and control policies in Canada, 1870-1970, pp. 88-ll6. In: J. Blackwell and P. Erickson (eds.), *Illicit Drugs in Canada: A Risky Business.* Toronto: Nelson Canada.

Solomon, R., E. Single and P. Erickson (1983). Legal considerations in Canadian cannabis policy. *Canadian Public Policy*, 9(4):419-433.

Stephenson, G. H., L. P. Lingley, G. E. Trasov and H. Stanfield (1956). *Drug Addiction in British Columbia.* Vancouver: University of British Columbia (unpublished manuscript).

Tokatlian, J. G. (1988). National security and drugs: Their impact on Colombian-U. S. relations. *Journal of Interamerican Studies and World Affairs*, 30(1):133-166.

Wagstaff, A. and A. Maynard (1988). *Economic Aspects of the Illicit Drug Market and Drug Enforcement Policies in the United Kingdom.* Home Office Research Unit Study, No. 95. London: HMSO.

Waley, A. (1958). *The Opium War Through Chinese Eyes.* Stanford: Stanford University Press.

The White House (1989). *National Drug Control Strategy.* Washington, D. C.: U.S. Government Printing Office.

Wisotsky, S. (1986). *Breaking the Impasse in the War on Drugs.* Westport, CT: Greenwood Press.

Zimring, F. E. and G. Hawkins (1973). *Deterrence: The Legal Threat in Crime Control.* Chicago: Univesity of Chicago Press.

40TH ANNIVERSARY SCIENTIFIC LECTURE SERIES

WINDOWS ON SCIENCE

Forty Years of Basic Research: Investment in the Future

HAROLD KALANT

ADDICTION RESEARCH FOUNDATION

TORONTO, ONTARIO

HAROLD KALANT

Harold Kalant received his M.D. from the University of Toronto in 1945. After several years of clinical work, he turned to basic research and obtained his PhD in 1955. In 1959 he was named Assistant Research Director, Biological Sciences, at the ARF. He later became Associate Research Director before receiving Emeritus status in 1989. With the late R. J. Gibbins and other colleagues, he explored the interactions between biological mechanisms and behavioral and environmental influences in the production of behavioral changes and tolerance by alcohol and other psychoactive drugs. His numerous awards include the Jellinek Award (1972), the Annual Research Award of the Research Society on Alcoholism (1983), the Upjohn Award of the Pharmacological Society of Canada (1985) and the Nathan B. Eddy Gold Medal of the Committee on Problems of Drug Dependence (1986). Dr. Kalant is President of the International Society for Biomedical Research on Alcoholism.

Previous lecturers in this Addiction Research Foundation 40th Anniversary series have covered, with great expertise and elegance, the evolution of scientific thought on specific topics in the addiction field over the past 40 years. The purpose of this final lecture is somewhat different: it is to show how basic research, carried out in this Foundation and elsewhere over relatively long periods of time, without reference to immediate specific applications, can alter knowledge and concepts in important ways and, in so doing, lead to unanticipated practical applications that are of very great value. To begin, however, it is necessary to make clear what "research" means, and specifically what basic research is, and is not.

WHAT IS BASIC RESEARCH ?

In popular usage, research is widely understood to mean searching out and assembling existing information about a subject from any available sources, for the purpose of presentation, clarification, or decision-making. This is research as practised by a researcher for a television program, an investigative journalist, a researcher for a government ministry, or a student preparing an essay. To the scientist, however, research consists of the systematic, logical investigation of a phenomenon, which generates *new* knowledge that helps to explain that and other related phenomena, and that permits testable predictions of their consequences.

The difference between *basic* and *applied* research is not in their methods and governing principles, but in their purposes. Basic research is conducted to increase our knowledge and understanding of a phenomenon or process, its nature, cause(s), mechanism(s), modifying or controlling factors, and consequences; the knowledge and understanding are themselves the primary goals. Applied research is carried out to discover how the knowledge can be applied to the solution of a specific practical problem. But to the researchers engaged in these activities, no sharp boundary exists. A basic biochemist studying the molecular biology of DNA in tumor cells in mice will usually not refuse to look at the potential application of his discoveries to the treatment of cancer in humans. Conversely, a physical chemist, working in the laboratories of a major pharmaceutical company to develop a new or improved drug, may have to do basic research on the molecular configurations of existing drugs in order to know where improvements may be made. It is extremely important to emphasize that basic and applied research are on a continuum, and the differences between them are only matters of degree.

The nature of basic research is independent of the specific discipline in which it is carried out. It was recently described rather scornfully, by some one quite unfamiliar with its actual operation, as "playing with rats in the basement."

This is a total misunderstanding of basic research. It may deal with biomedical, psychological, sociological, economic, physical, chemical, astronomical or any other type of subject matter, but it is defined only by its purpose of generating new knowledge.

To illustrate how wide the range of basic research is, a few examples from within ARF may suffice: the work of Carlen *et al.* on brain imaging in alcoholics (1); the studies by Popham and Schmidt on the effects of price on the distribution-of-consumption curve for alcohol (2); experimental exploration by Corrigall, Linseman and Mucha of the brain mechanisms of reinforcement of self-administration of alcohol or opiates by rats (3-5); and the examination by Sellers and Naranjo of the effects of serotonin reuptake inhibitors on ethanol drinking by both rats and humans (6, 7). At the Addiction Research Foundation, and in the addiction research field in general, the knowledge generated by all such research is ultimately intended to lead to improved methods of preventing, detecting, or treating addiction and its consequences. Thus, in this field all basic research *is consciously intended to result in future applied research.*

WHERE HAVE WE COME FROM ?

In order to assess the progress made during the past 40 years, it is useful to look back at what was known about alcohol and alcoholism in 1949. At the risk of considerable oversimplification, we can summarize it as follows:

- About the beginning of this century, Meyer and Overton demonstrated that alcohols, like volatile anesthetics, distributed or partitioned themselves between fatty materials and water, to degrees that were characteristic of the individual alcohols or anesthetics. Moreover, they affected cell functions in nervous and other tissues in direct proportion to their partition coefficients. This led to the hypothesis that they dissolved in the fatty materials that make up the bulk of the outer membranes surrounding all cells, altered the properties of the membranes, and thus affected cell functions (8). Over the next half-century, great advances were made in characterizing these effects on the electrophysiological and biochemical properties of nerve cell membranes. By 1949, however, all such effects were still demonstrable only at very high alcohol concentrations that would be profoundly intoxicating or lethal in the living animal. Therefore the mechanism of action of alcohol on living cells, which gave rise to mild or moderate degrees of intoxication, was still unknown.

- It was known that ethanol was oxidized, primarily by alcohol dehydrogenase in the

liver, to acetaldehyde. This highly toxic substance was also known to be rapidly oxidized by acetaldehyde dehydrogenase to acetate, and ultimately to carbon dioxide and water. However, the exact relationship of this whole process to general metabolism was not known (9). The oxidation of alcohol was believed to generate "empty calories," and its main consequences were believed to be due to the fact that heavy drinkers usually had a reduced intake of other nutrients normally found in a well balanced diet.

• The behavior patterns seen in alcoholics were well known, but argument still raged as to whether alcoholism was a 'disease' or merely badly-learned and maladaptive behavior (10). One may note, in passing, that there was not then, and still is not now, any agreed definition of what was meant by alcoholism as a 'disease.'

• Alcoholism was known to run in families, and in more recent work it was shown that the intensity of family history of alcoholism in the generations of the proband, the parents and the grandparents is an excellent predictor of the rate of progression of the problem in the proband (11). But there was no sound evidence available to settle the dispute about whether this familial pattern reflected genetic transmission of the problem from parents to children, or the influence of an alcoholic family environment on the children.

• Most of the characteristic organ pathologies caused by heavy drinking were known, including such examples as liver cirrhosis, various types of brain damage, peripheral neuropathy, alcoholic gastritis, and pancreatitis (12). It was not known, however, whether these were caused by direct toxic action of alcohol, by malnutrition secondary to excessive drinking, by indirect effects due to the metabolism of ethanol or by combinations of these factors and others.

• The existence of tolerance and physical dependence had been clearly recognized, but it was not clear to what extent tolerance was due to an increased rate of metabolic elimination of alcohol, and to what extent it reflected adaptive changes in the brain (13). Delirium tremens ("the DTs") had long been known clinically, but it was still not clear whether it was a form of pathological intoxication or a withdrawal reaction. The clear demonstration that it is a withdrawal reaction did not come until the outstanding work of the Lexington group in the mid-1950s (14).

• The magnitude of alcoholism as a public health problem was unknown. Jellinek had developed his mathematical

formula for estimating the prevalence of alcoholism in a population from the death rate from alcoholic cirrhosis (15), but the validity of his formula was disputed, as well as its degree of applicability in different societies.

HOW FAR HAVE WE COME ?

In order to trace how far our knowledge has advanced over the past 40 years, we can look at how our understanding of some of the foregoing questions has changed.

Mechanisms of action on cells

Interest in the Meyer-Overton hypothesis continued for many years, and even in the 1970s it was proposed that ethanol and other alcohols exerted their basic actions by partitioning in the lipid (fatty) layers of the cell membrane and disturbing their normal molecular architecture. This disturbance was shown to give rise to a loosening of intermolecular bonds, thus "fluidizing" the membrane (16,17). The fluidization was postulated to disrupt ion channels through which occurred the movements of sodium, potassium and other ions that underlie the excitation and inhibition of normally functioning nerve and muscle cells.

However, these effects still occurred only at very high alcohol concentrations. Throughout the 1960s and 1970s, neuroscientists therefore continued to examine the effects of alcohols on more specific cell functions, such as the activities of various constituent enzymes located in cell membranes, release of all the known neurotransmitters (chemical messengers between nerve cells), the binding of these transmitters to their specific receptors on the surfaces of their target cells, and the chemical signalling systems ("second messengers") that conveyed the effects of transmitter-receptor interaction across the membrane into the cell interior where responses occurred. All these processes, however, also proved to be resistant to the direct effects of ethanol except at very high concentrations (18-21).

In the late 1980s, major technical advances of several types suddenly changed the picture dramatically. One such advance was the development of the patch-clamp technique, which permitted the measurement of drug effects on a single type of ion channel in the cell membrane, or even on one single channel (22). Another was the development of methods for preparing suspensions of microscopic fragments of cell membranes, which retained the ability to respond to neurotransmitters by altering the movements of different ions across the membrane (23). Using these techniques, several groups of researchers were able to identify specific actions of alcohol on some of these receptor-linked channels (especially the sodium and calcium channel linked to the NMDA type of

glutamate receptor) at alcohol concentrations as low as 25-50 mg per 100 ml, well below the legal limit for driving an automobile (24, 25). Thus, for the first time there are now highly promising points of attack for elucidating the basic cellular actions of alcohol, including those responsible for its rewarding effects, for the development of tolerance, and possibly for some types of alcohol-related brain damage.

Alcohol metabolism

In the early 1960s, Lieber, Isselbacher, Lundquist and many others made important advances in clarifying the manner in which the metabolism of alcohol affected many important metabolic pathways in the liver and other organs (9, 26). During the oxidation of ethanol, the enzyme alcohol dehydrogenase transfers hydrogen from the alcohol to the coenzyme NAD, converting it to NADH. The increased concentration of NADH interferes with the oxidation of fatty acids, glycerol, and numerous other substances by enzymes that also use NAD as a co-factor. One result is that glycerol and fatty acids accumulate, and combine to form neutral fat (triglycerides) in larger amounts than normal, so that fat droplets rapidly accumulate in the liver. Lieber *et al.* showed convincingly that one weekend of heavy drinking could produce a fatty liver, even in healthy normal humans and rats consuming a nutritious, well-balanced diet (27).

It had been known since 1935 that other enzymes, such as catalase, could play a minor role in the oxidation of ethanol (28), but no great biological importance was attached to this fact. In the late 1960s, however, Lieber and De Carli reported that an enzyme system contained in liver microsomes could also oxidize ethanol. They named this system "microsomal ethanol oxidizing system" or MEOS, and showed that it increased in amount in the livers of rats and humans who regularly consumed large quantities of alcohol (29). For nearly two decades, there was a heated debate about the nature of MEOS and its claimed role *in vivo* (9). The debate was finally resolved in the late 1980s, when modern molecular biochemical techniques using monoclonal antibodies demonstrated conclusively that MEOS does indeed exist in the normal liver, and that its central component is a specific enzyme dubbed cytochrome P450IIE1 (30). Though this enzyme plays only a minor role in alcohol metabolism in moderate drinkers, it increases greatly in amount in heavy drinkers and may play an important role in metabolic tolerance to alcohol. More importantly, perhaps, it also metabolizes a number of other drugs, and the increase of this enzyme in the livers of heavy drinkers produces metabolic cross-tolerance to these drugs (31). It will now be possible for clinical pharmacologists to predict with considerable accuracy the metabolic interactions between alcohol and new drugs that are metabolized by the cytochrome system in the liver.

The role of acetaldehyde dehydrogenase had been exploited, since the 1940s and early 1950s, in the clinical treatment of alcoholics. First disulfiram (Antabuse®), and later citrated calcium carbimide (Temposil®), were found to inhibit acetaldehyde dehydrogenase, so that acetaldehyde would accumulate in the blood if the person drank alcohol, and this would produce a violently unpleasant reaction that was intended to discourage the person from drinking alcohol again (32, 33).

In the 1970s, however, the enzyme attracted attention for a different reason. Among Oriental peoples, about half of those who try alcohol react to even small amounts of it with intense flushing, tachycardia, nausea and a sharp fall in blood pressure, which mirror closely the reaction provoked by disulfiram. As a result, alcohol use and alcoholism are very much less frequent among those who react in this way. Since this reaction was found to occur even in Oriental infants to whom alcohol was administered (34), it seemed very likely that the reaction was genetically caused. The flushing reaction was later found to be caused by a genetic difference in the properties of the acetaldehyde dehydrogenase present in the 50 per cent of Orientals who flush when they drink alcohol. They have a variant of the enzyme that has very low efficiency in oxidizing acetaldehyde, whereas the 50 per cent who do not flush with alcohol have the normal form of the enzyme (35). Thus, possession of the inefficient form constitutes an inborn protective factor that diminishes the likelihood of drinking alcohol, and of becoming alcoholic (36, 37).

Acetaldehyde itself is a highly reactive substance, that combines rapidly and non-enzymatically with amines and other substances. Much attention has been devoted in recent years to the possibility that some of the effects previously attributed to alcohol are in fact due to acetaldehyde combining with amine groups in a variety of biologically important molecules. In the 1970s, Peterson and others began to examine the possibility that acetaldehyde could combine with hemoglobin in the circulating blood, to form a stable adduct that might serve as an indicator of the amount of alcohol the person had consumed in the previous few months (38). This concept has been extended by the work of a number of other research groups (39-41), including that of Yedy Israel at ARF (42), which have shown that acetaldehyde forms stable adducts with a large variety of plasma and tissue proteins. These adducts can act as antigens, giving rise to the production of specific antibodies by the immune system. The antibodies can be measured in the circulating blood, or they can be used as a tool for measuring the acetaldehyde-protein adducts in the blood by immunoassays of various types. Both antibody levels and adduct levels are being tested as possible methods for estimating the amount of alcohol consumption during the preceding period corresponding to the lifetime of the proteins in question (43,

44). Such methods offer the possibility of development of a relatively inexpensive, objective screening test for identifying people who drink excessively, before they actually produce physical harm to themselves. In addition, current research is exploring the possibility that alcohol-induced cell damage in various organs may be initiated by the formation of acetaldehyde adducts with essential structural proteins such as intracellular microtubules (45), or by reactions between the adducts and some of the antibodies on the surface of liver cells (46).

Liver pathology

Apart from the possible role of acetaldehyde as noted above, liver cell damage may occur in heavy drinkers for a different reason. Beginning in the early 1960s, work by ARF researchers showed that chronic ingestion of large amounts of alcohol led to an increase in the activity of a cell membrane enzyme, sodium-potassium-ATPase, in the liver, brain and other organs (47). This enzyme splits the very important regulator substance ATP (adenosine triphosphate), and thus stimulates the metabolic utilization of oxygen in a series of reactions that synthesize more ATP to replace that which has been split. In the early 1970s, Israel proposed that in the liver, which normally receives a relatively lower oxygen supply than most other tissues, the extra oxygen consumption caused by the increased ATPase activity could result in serious oxygen deficiency in those liver cells farthest from the incoming blood supply (48). According to this hypothesis, alcohol-induced liver cell damage might really be hypoxic damage, and it would be made worse by anything that reduced the oxygen supply further, and improved by anything that reduced the oxygen demand in liver cells.

This hypothesis led to the testing of the antithyroid compound propylthiouracil (PTU), which was known to reduce the hepatic oxygen uptake, as a treatment for acute alcoholic hepatitis. After animal experiments had demonstrated that PTU did indeed reduce the vulnerability of the liver to hypoxic damage in rats receiving high-alcohol diets (48), a short-term test was conducted in patients with alcoholic liver disease in the Clinical Institute of the ARF. PTU accelerated the recovery of liver function, so that normalization occurred in only half of the time required by patients not receiving PTU (49). This success led to a long-term trial in which PTU was given to patients as required by the state of their liver function, for up to two years (50). The results were striking. In patients with proven active alcoholic hepatitis who received a treatment program that included PTU, the cumulative death rate over the period of the study was cut by more than 60 per cent, compared to others receiving the same treatment program without PTU. This is the first treatment that has succeeded in cutting significantly the death rate from alcoholic hepatitis.

Genetic influences in alcoholism

This topic was covered in detail by Dr. Cloninger in an earlier lecture in this series (*see Chapter 3*), and will not be examined in detail here. At this point it is sufficient to recall a few of the main types of evidence, acquired during the 1960s and later, that demonstrated the importance of genetic factors in determining individual differences in susceptibility to alcoholism or alcohol-related problems (51).

Human studies showed that if an alcoholic patient had a twin, the concordance for alcohol problems in the twin was higher in monozygotic (identical) twins than in dizygotic (fraternal) twins. This finding was generally interpreted as evidence of a genetic factor, although it was recognized that monozygotic twins are likely to have more nearly identical environments than dizygotic twins or ordinary siblings do. Further support for a genetic interpretation came from the finding that the sons of alcoholic biological fathers, put out for adoption by unrelated adoptive parents before the age of three months, had a much higher likelihood of developing alcohol problems as adults than similarly adopted sons of non-alcoholic biological fathers did. This finding has been confirmed in several subsequent studies in other countries.

Animal experiments lent further support to the view that genetic factors play an important role. In the late 1950s and early 1960s, Mardones *et al.* described the genetic selection of two diverging lines of rats (UChA and UChB) from a common starting stock, on the basis of their high and low voluntary consumption of alcohol solutions (52). Similar genetic selections of high and low drinking lines were described in the 1960s in Finland (AA and ANA lines) and later in the USA (P and NP, HAD and LAD lines) (53, 54). Since then, there has been a rapid expansion of such work, with genetic selections for low and high sensitivity to the effects of ethanol, low and high proneness to alcohol withdrawal seizures, and so forth. It is not surprising that such genetic influences have been demonstrable. Virtually every action of ethanol in the body, and every metabolic action of the body on ethanol, is dependent on a protein of some sort — an enzyme, a membrane constituent, a receptor, a structural protein, etc. — and the synthesis of each protein is regulated by a gene. Therefore there is a high possibility of genetic variation in a great many of these interactions between the alcohol and the biological system.

The question therefore becomes one of finding out exactly what is inherited that results in the genetically-determined differences in susceptibility to high alcohol intake and its consequences. With the advent of molecular biology in the 1980s, the techniques have rapidly become available for identifying protein polymorphisms, comparing the presence of the different forms of specific proteins in drinkers versus non-drinkers or alcohol-sensitive versus alcohol-resistant lines, and tracing these protein differences back

to the nucleic acids that produce them. By identifying the specific products of these genes, it should then be possible to identify the exact biological roles of these products and thus determine how they contribute to the differences in response to alcohol (55).

The differences could be of many kinds. There could be differences in taste or smell that make alcohol more attractive or more repulsive. The reinforcement or 'reward' mechanisms in the brain might be more sensitive to alcohol, or the aversive or 'punishment' mechanisms might be less sensitive. The high drinker might be able to develop nervous system tolerance more rapidly than the low drinker, and therefore be able to withstand more alcohol. The low drinker might be more sensitive to alcohol-induced damage to the stomach lining, and therefore be forced to stop drinking sooner. The possibilities are almost endless.

One example from ARF research is provided by the work of Grupp and his colleagues, who found that a wide variety of factors, having in common their ability to increase activity of the renin-angiotensin system in the brain of the rat, would decrease the rat's voluntary consumption of alcohol. Factors that decreased the renin-angiotensin activity caused an increase in alcohol intake (56). Examination of the P (alcohol-preferring) and NP (non-preferring) lines revealed low renin levels in the P rats, suggesting that the genetic difference might be one that results in a deficient production of renin or angiotensin in the heavy-drinking P line (57). Very recently, implantation of angiotensin-producing fetal cells into the third ventricle of the brain was found to reduce alcohol intake (A. J. Lança and L. A. Grupp, unpublished findings). With the development of more techniques of genetic engineering, it may be possible to remedy some of the genetic alterations found to give rise to heavy drinking or to greater sensitivity to alcohol-induced damage.

It is important to remember, however, that these genetic differences do not directly produce alcoholism. They produce greater or lesser sensitivity to specific effects of alcohol, which may increase or decrease the *risk* of becoming alcoholic if one drinks. Yet it is well known that children of alcoholics, bearing the putative genetic factors, may choose not to drink, and never become alcoholics. In the scientific search for genes and their products, one must not lose sight of the fact that social and psychological factors play a very important role, and that it is the interaction between them and the biological factors that determines the final outcome.

Tolerance

In the 1950s, McQuarrie and Fingl (58) demonstrated clearly that alcohol tolerance could result from changes in sensitivity of the brain to a given blood concentration of alcohol. Thus, tolerance must represent adaptive changes in brain cells, and not

merely an increased rate of removal of alcohol from the blood by the liver and other organs. This was confirmed later by the demonstration that tolerance involved a reduced effect of a given concentration of alcohol in the brain itself (59). This was generally interpreted as a homeostatic response to the presence of the drug itself in the brain. A technique for producing alcohol tolerance in mice or rats, introduced in the 1970s, consisted of placing the animals in a closed chamber in which the air that they breathed contained a fixed concentration of alcohol vapor; after a few days, they showed tolerance and physical dependence (60). This appeared to support the concept that the mere presence of the drug in the brain was the effective stimulus to the development of tolerance.

In 1968, however, Chen showed that tolerance developed rapidly in rats that had to practise a task every day under the influence of a small dose of alcohol, whereas other rats that received the same daily dose *after* practising the task did not become tolerant (61). This stimulated a long series of studies by ARF researchers, which led to the concept that tolerance develops in response to the functional disturbance produced by alcohol (or other drugs), rather than to the mere presence of the drug itself (62). Repeated performance of a task under the influence of the drug (63, 64), or Pavlovian conditional linkage of drug administration to specific environmental cues (that thus acquired the ability to predict drug administration), were found to enhance markedly the degree or the speed of appearance of tolerance (65). If these environmental cues were then presented but were not followed by drug administration, they provoked withdrawal reactions. All of these studies revealed that learning played a major role in the production of tolerance and physical dependence, and that it could therefore, in principle, play an important role in reversing them.

Concurrent neurochemical studies, by ARF researchers and others, demonstrated that a number of neuronal systems and pathways known to play important roles in processes of learning and memory were also important in tolerance. For example, administration of inhibitors of cerebral protein synthesis, which had been shown to prevent the learning of new skills without interfering with the performance of previously learned ones, also prevented the acquisition of tolerance to alcohol (66, 67). Administration of vasopressin or certain related peptides — which had been shown previously to maintain a learned behavioral response under conditions in which it would have been lost without the peptides (68) — was also found to retain alcohol tolerance under conditions in which it would otherwise have been lost — i.e., after the suspension of alcohol intake (69, 70). This tolerance-maintaining effect of vasopressin was found to occur only if a specific group of nerve fibres from the median raphe nucleus of the brainstem, which releases serotonin at specific sites in the hippocampus and septum, were intact.

The exact site and nature of the interaction between serotonin and vasopressin that underlies the retention of learned behaviors and of alcohol tolerance are now under active investigation. The possible future implications of all this work are very great. If the ability to develop tolerance is one important determinant of the level of alcohol consumption, as many researchers believe, then the discovery of all the behavioral, environmental and neurochemical mechanisms that contribute to the development and maintenance of tolerance will open many possibilities of therapeutic intervention for decreasing tolerance and thus reducing alcohol intake.

Epidemiological studies

In the early 1960s, the Jellinek formula for estimating the prevalence of alcoholism, which had continued to be the subject of uncertainty, was validated in a thorough study by Popham (71, 72). Subsequently, the ARF epidemiology research group showed, in multi-country comparisons at one point in time, and in single-country comparisons over time, that there was a very high positive correlation between the mean per capita consumption of alcohol and the death rate from alcoholic cirrhosis (73). At the same time, time-trend studies prompted by the work of an excellent investigative journalist with the Toronto *Globe & Mail*, Terry Greer, established that the mean price of beverage alcohol in constant dollars was inversely correlated with the mean per capita consumption and with the death rate from alcoholic cirrhosis (2). This inverse correlation was later confirmed experimentally in an elegant clinical research project that was specifically prompted by the ARF epidemiological work (74).

In 1956, the French epidemiologist Ledermann proposed a model of the distribution of different levels of alcohol consumption in a population, which could be described by a unimodal curve, skewed toward the upper end of the consumption scale (75). This model predicted that the majority of drinkers would be clustered around a modal value relatively close to the lower end of the scale, with steadily decreasing numbers of users at progressively higher levels of intake. Ledermann's hypothesis predicted that any change in the mean per capita consumption by the whole population would be accompanied by a corresponding shift of the modal value, and corresponding changes in the numbers of users at the high end of the scale, i.e., those at risk of alcohol-related problems. The ARF group found that this hypothesis was borne out by the data from Canada and several other countries, any shift in the modal level of consumption being reflected in a corresponding change in the death rate from alcoholic cirrhosis (73). Smart *et al.* later showed similar curves of distribution of consumption for other psychoactive drugs among the population of students in Ontario (76).

These findings have obvious implications for public policies on control

of the sale and use of alcohol and other drugs. If a major reduction in the harmful consequences of heavy use can be achieved by reduction of the mean per capita use by the whole population, and if an increase in price is an effective means of achieving such a reduction, this suggests that a social experiment involving increases in price or in taxation of alcohol might be the most cost-effective means of lowering the harmful effects of alcoholism on health and social functioning.

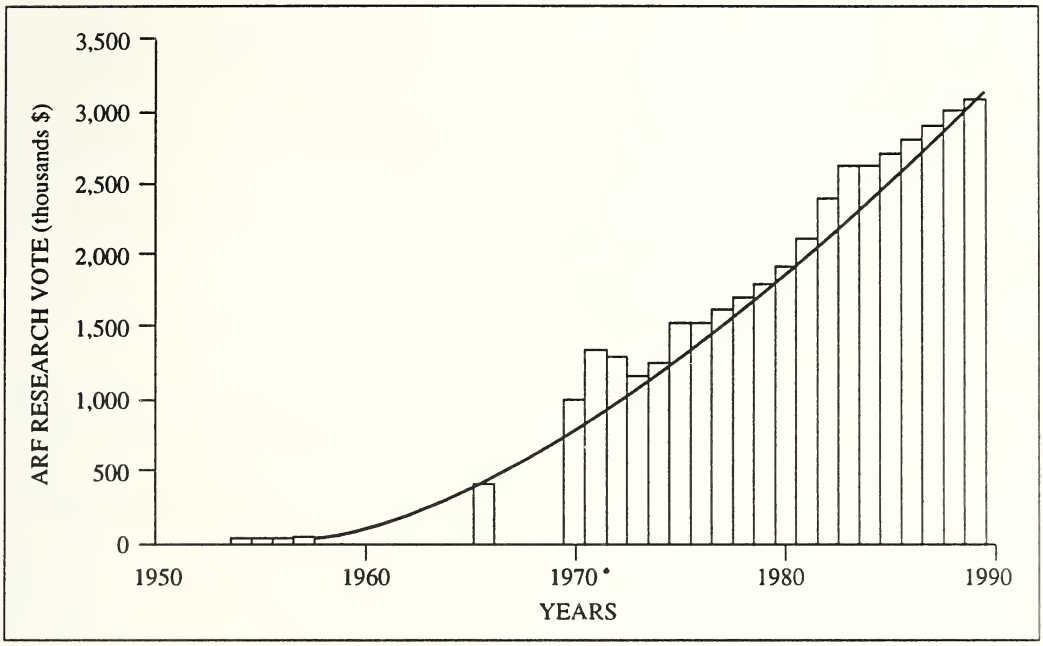
LESSONS FROM THE HISTORY OF ALCOHOL RESEARCH

The foregoing review of advances in alcohol research over the past 40 years shows that there is often a long developmental lag time between the start of research on a particular question and the time that the knowledge gained from that research can be practically applied. For example, the demonstration that PTU has a clear beneficial effect on the course of alcoholic liver disease came only after about 25 years of research. The concept of price as a potentially effective controlling factor, to reduce the prevalence and costs of alcoholism evolved from more than 20 years of continuous methodical research on the same line.

Governments and citizens alike may therefore ask, *Does it pay?* Fortunately, the answer is relatively easy to provide. Figure 1 illustrates the growth of ARF basic research expenditures, as gleaned from data in the annual reports from the early 1950s up to 1989. For comparison, the annual statistics of the Ontario health care system for 1984/85 indicate that there were about 3,335 admissions a year for alcoholic liver disease, and the average duration of stay was 15.6 days, at an average daily cost of $318.18 per bed. Thus, the bed costs alone were nearly $18 million for those patients, leaving aside the costs of special diagnostic and therapeutic procedures, lost work time, and social welfare. The first clinical study on PTU indicated that it cut the recovery time by 50 per cent. Therefore, if applied throughout the health care system, this treatment would have reduced the hospital bed costs by about $9 million that year, or more than three times the entire basic research budget of the Foundation for that year. On that basis, the hospital savings resulting from that one research program alone would, over the following 10 years, more than repay the entire basic research expenditures of the Foundation since its inception. If one adds to this the economies that might result from other unrelated lines of research, such as the potentially huge monetary savings that could result from the adoption of alcohol price policies suggested by the epidemiological work, the cost of basic research would be recouped many times over in a very short time, even without taking into account the immense non-monetary benefits.

Successful major industrial

Figure 1. Annual budgets for the Research Division (and its predecessors and successors) of the Addiction Research Foundation of Ontario, as reported in the Foundation's Annual Reports for the years shown. No information was available for the years corresponding to the missing bars. **Note:** The actual amounts shown should be regarded as approximations, because the Foundation changed its method of reporting research expenditures several times during the period shown, and exact correspondence of the figures for different years cannot be assumed.



corporations have long known that research pays. For example, Bell Telephone laboratories supported Karl Jansky's pioneering research on radio-astronomy for nearly a decade without any early practical application in mind (77). When the need for greatly improved radio-communications became important during World War II, and work on radar development began, the practical applications of Jansky's discoveries rapidly appeared and the costs of his research program were quickly justified. Basic research in radio astronomy has continued in the Bell laboratories right up to the present, and in fact earned a Nobel prize for Wilson and Penzias (78) in 1965. Progressive corporations know that without today's basic research there will be no applications to develop in the longer-term future (79). Unfortunately, Canadian and American firms, on the whole, have not yet learned this as well as their survival requires. In 1988, for example, R & D expenditures by General Motors and IBM were in the range of four to five *billion* dollars a year. Yet even these huge expenditures are outdone by the amounts spent by Japanese and European competitors. There is an excellent correlation between

R & D expenditure, as a percentage of corporate revenue, and the growth in a corporation's market share during the following 10 years (80). The steady erosion of North American industrial predominance is attributed in large part to the 50 per cent greater research expenditures by more successful European and Japanese competitors (81).

Governments, research institutions, and the general public must be made aware that "playing with rats in the basement" or "playing with statistics in a library" are not idle amusements for the researcher at public expense, but sound business sense. In addition to their intrinsic intellectual and human value, they generate new and more effective ways of improving health and socio-economic functioning, and at lower cost. Basic research is indeed an investment in the future.

REFERENCES

1. Carlen PL, Penn RD, Fornazzari L, Bennett J, Wilkinson DA and Wortzman G (1986). Computerized tomographic scan assessment of alcoholic brain damage and its potential reversibility. Alcoholism Clin. Exp. Res. 10: 226-232.

2. Bruun K, Edwards G, Lumio M, Mäkelä K, Tan L, Popham RE, Room R, Schmidt W, Skog OJ Sulkunen P and Østerberg E (1975). Alcohol Control Policies in Public Health Perspective. Finnish Foundation for Alcohol Studies, vol. 25. Finnish Foundation, Helsinki.

3. Corrigall WA and Linseman MA (1988). Conditioned place preference produced by intra-hippocampal morphine. Pharmacol. Biochem. Behav. 30: 787-789.

4. Linseman MA and Harding S (1990). Intracerebroventricular morphine enhances alcohol consumption by rats. Pharmacol. Biochem. Behav. 36: 405-408.

5. Mucha RF, Millan MJ and Herz A (1985). Aversive properties of naloxone in non-dependent (naive) rats may involve blockade of central b-endorphin. Psychopharmacology 86: 281-285.

6. Naranjo CA, Sullivan JT, Lawren MO and Sellers EM (1987). Strategies for the identification and testing of new pharmacological modulators of ethanol consumption. In J Engel and L Oreland (eds.), Brain Reward Systems and Abuse. New York: Raven Press. pp. 129-145.

8. Seeman P (1972). The membrane actions of anesthetics and tranquilizers. Pharmacol. Rev. 24: 583-655.

9. Kalant H (1991). Research on alcohol metabolism: a historical perspective. Keio J Med 40: 113-117.

10. Jellinek EM (1960). The Disease Concept of Alcoholism. New Haven: Hillhouse Press.

11. Volicer L, Volicer BJ and D'Angelo N (1985). Assessment of genetic predisposition to alcoholism in male alcoholics. Alc. Alcoholism 20: 63-68.

12. Kissin B and Begleiter H, eds. (1974). The Biology of Alcoholism, vol. 3: Clinical Pathology. New York: Plenum Press.

13. Newman HW, Lehman AJ (1937). Nature of acquired tolerance to alcohol. J. Pharmacol. Exp. Ther. 61: 301-305.

14. Isbell H, Fraser HF, Wikler A, Belleville RE and Eisenman AJ (1955). An experimental study of the etiology of "rum fits" and delirium tremens. Quart. J. Stud. Alcohol 16: 1-33.

15. Jellinek EM (1947). Recent trends in alcoholism and in alcohol consumption. Quart. J. Stud. Alcohol 8: 1-42.

17. Chin JH and Goldstein DB (1977). Drug tolerance in biomembranes: a spin label study of the effects of ethanol. Science 196: 684-685.

18. Kalant H and Woo N (1981). Electrophysiological effects of ethanol on the nervous system. Pharmacol. Ther. 14: 431-457.

19. Tabakoff B and Hoffman PL (1980). Alcohol and neurotransmitters. In Rigter H and Crabbe JC Jr (eds.), Alcohol Tolerance and Dependence. Amsterdam: Elsevier. pp. 201-226.

20. Tabakoff B, Hoffman PL and Liljequist S (1987). Effects of ethanol on the activity of brain enzymes. Enzyme 37: 70-86.

21. Deitrich RA, Dunwiddie TV, Harris RA and Erwin VG (1989). Mechanism of action of ethanol: initial central nervous system actions. Pharmacol. Rev. 41: 489-537.

22. Neher E and Sakmann B (1976). Single-channel currents recorded from membrane of denervated frog muscle fibres. Nature 260: 779-802.

23. Hollingsworth EB, McNeal ET. Burton JL, Williams RJ and Daly JW (1985). Biochemical characterization of a filtered synaptoneurosome preparation from guinea pig cerebral cortex: cyclic adenosine 3':5'-monophosphate-generating system, receptors and enzymes. J. Neurosci. 5: 2240-2253.

24. Lovinger DM, White G and Weight FF (1989). Ethanol inhibits NMDA-activated ion current in hippocampal neurons. Science 243: 1721-1724.

25. Hoffman PL, Rabe CS, Moses F and Tabakoff B (1989). N-methyl-D-aspartate receptors and ethanol: inhibition of calcium flux and cyclic GMP production. J. Neurochem. 52: 1937-1940.

26. Lieber CS (1988). Biochemical and molecular basis of alcohol-induced injury to liver and other tissues. N. Engl. J. Med. 319: 1639-1650.

27. Lieber CS, Jones DP and De Carli LM (1965). Effects of prolonged ethanol intake: production of fatty liver despite adequate diets. J. Clin. Invest. 44: 1009-1021.

28. Keilin D and Hartree EF (1935). Coupled oxidation of alcohol. Proc. Roy. Soc. Lond. [B] 119: 141-159.

29. Lieber CS and De Carli LM (1968). Ethanol oxidation by microsomes: adaptive increase after ethanol feeding. Science 162: 917-918.

30. Ekström G, von Bahr C and Ingelman-Sundberg M (1989). Human liver microsomal cytochrome P450 IIE1. Immunological evaluation of its contribution to microsomal ethanol oxidation, carbon tetrachloride reduction and NADPH oxidase activity. Biochem. Pharmacol. 38: 689-693.

31. Koop DR and Coon MJ (1986). Ethanol oxidation and toxicity: role of alcohol P-450 oxygenase. Alcoholism Clin. Exp. Res. 10: 44S-49S.

32. Hald J and Jacobsen E (1948). The formation of acetaldehyde in the organism after ingestion of Antabuse (tetraethylthiuramdisulphide) and alcohol. Acta Pharmacol. Toxicol. 4: 305-3.

33. Ferguson JKW (1956). A new drug for alcoholism treatment. I. A new drug for the treatment of alcoholism. Can. Med. Assoc. J.74: 793-795.

34. Wolff PH (1972). Ethnic differences in alcohol sensitivity. Science 175: 449-450.

35. Goedde HW, Agarwal DP, Eckey R and Harada S (1985). Population genetic and family studies on aldehyde dehydrogenase deficiency and alcohol sensitivity. Alcohol 2: 383-390.

36. Higuchi S, Muramatsu T, Shigemori K, Saito M, Kono H, Dufour MC and Harford TC (1992). The relationship between low Km aldehyde dehydrogenase phenotype and drinking behavior in Japanese. J. Studies Alc. (in press).

37. Ohmori T, Koyama T, Chen C-C, Yeh E-K, Reyes BV Jr. and Yamashita I (1986). The role of aldehyde dehydrogenase isozyme variance in alcohol sensitivity, drinking habits formation and the development of alcoholism in Japan, Taiwan and the Philippines. Progr. Neuropsychopharmacol. Biol. Psychiat. 10: 229-235.

38. Stevens VJ, Fantl WJ, Newman CB, Sims RV, Cerami A and Peterson CM (1981). Acetaldehyde adducts with hemoglobin. J. Clin. Invest. 67: 361-369.

39. Tuma DJ, Hoffman t and Sorrell MF (1991). The chemistry of acetaldehyde-protein adducts. In Kalant H, Khanna JM and Israel Y (eds.), Alcohol & Alcoholism, Suppl. 1:271-276.

40. Behrens UJ, Hoerner M, Lasker JM and Lieber CS (1988). Formation of acetaldehyde adducts with ethanol-inducible P450IIE1. Biochem. Biophys. Res. Commun. 154:584-590.

41. Lin RC, Smith RS and Lumeng L (1988). Detection of a protein-acetaldehyde adduct in the liver of rats fed ethanol chronically. Clin. Invest. 81:615-619.

42. Israel Y, Hurwitz E, Niemelä, O and Arnon R (1986). Monoclonal and polyclonal antibodies against acetaldehyde-containing epitopes in acetaldehyde-protein adducts. Proc. Natl. Acad. Sci. USA 83: 7923-7927.

43. Worrall S, De Jersey J, Shanley BC and Wilce PA (1991). Alcohol abusers exhibit a higher IgA response to acetaldehyde-modified proteins. In Kalant H, Khanna JM and Israel Y (eds.), Alcohol & Alcoholism, Suppl 1: 261-264.

44. Niemelä O, Klajner F, Orrego H, Vidins E, Blendis LM and Israel Y (1987). Antibodies against acetaldehyde-modified protein epitopes in human alcoholics. Hepatology 7: 1210-1214.

45. Tuma DJ, Jennett RB and Sorrell MF (1987). The interaction of acetaldehyde with tubulin. Ann. N.Y. Acad. Sci. 492:277-286.

46. Sorrell MF and Tuma DJ (1985). Hypothesis: Alcoholic liver injury and the covalent binding of acetaldehyde. Alcoholism: Clin. Exp. Res. 9: 306-309.

47. Kalant H (1985). Tolerance, learning, and neurochemical adaptation. Can. J. Physiol. Pharmacol. 63: 1485-1494.

48. Israel Y, Kalant H, Orrego H, Khanna JM, Phillips MJ and Stewart DJ (1979). Hypermetabolic state, oxygen availability, and alcohol-induced liver damage. In Majchrowicz E and Noble EP (eds.), Biochemistry and Pharmacology of Ethanol, Vol. 1. New York: Plenum Publishing. pp. 433-444.

49. Orrego H, Kalant H, Israel Y, Blake J, Medline A, Rankin JG, Armstrong A and Kapur B (1979). Effect of short-term therapy with propylthiouracil in patients with alcoholic liver disease. Gastroenterology 76: 105-115.

50. Orrego H, Blake JE, Blendis LM, Compton KV and Israel Y (1987). Long-term treatment of alcoholic liver disease with propylthiouracil. New Engl. J. Med. 317: 1421-1427.

51. Schuckit MA, Li T-K, Cloninger CR and Deitrich RA (1985). Genetics of alcoholism. Alcoholism Clin. Exp. Res. 9: 475-492.

52. Mardones J and Segovia-Riquelme N (1983). Thirty-two years of selection of rats by ethanol preference: UChA and UChB strains. Neurobehav. Toxicol. Teratol. 5: 171-178.

53. Eriksson K (1968). Genetic selection for voluntary alcohol consumption in the albino rat. Science 159:739-741.

54. Froehlich JC and Li T-K (1991). Animal models for the study of alcoholism. J. Addict. Dis. 10: 61-71.

55. Worton RG (1991). Molecular genetic approaches to the study of individual risk in alcoholism. In Kalant H, Khanna JM and Israel Y (eds.), Alcohol & Alcoholism, Suppl. 1: 19-25.

56. Grupp LA, Perlanski E and Stewart RB (1991). Regulation of alcohol consumption by the renin-angiotensin system: a review of recent findings and a possible mechanism of action. Neurosci. Biobehav. Rev. 15: 265-275.

57. Grupp LA, Kalant H and Leenen FHH (1989). Alcohol intake is inversely related to plasma renin activity in the genetically selected alcohol-preferring and -nonpreferring lines of rats. Pharmacol. Biochem. Behav. 32: 1061-1063.

58. McQuarrie DG and Fingl E (1958). Effects of single doses and chronic administration of ethanol on experimental seizures in mice. J. Pharmacol. Exp. Ther. 124: 264-271.

59. LeBlanc AE, Kalant H and Gibbins RJ (1975). Acute tolerance to ethanol in the rat. Psychopharmacologia (Berl) 41: 43-46.

60. Goldstein DB and Pal N (1971). Alcohol dependence produced in mice by inhalation of ethanol: grading the withdrawal reaction. Science 172: 288-290.

61. Chen C-S (1968). A study of the alcohol-tolerance effect and an introduction of a new behavioral technique. Psychopharmacologia (Berl) 12: 433-440.

62. Kalant H, LeBlanc AE and Gibbins RJ (1971). Tolerance to, and dependence on, some non-opiate psychotropic drugs. Pharmacol. Rev. 23: 135-191.

63. Lê AD, Kalant H and Khanna JM (1987). Effects of treatment dose and intoxicated practice on the development of tolerance to ethanol-induced motor impairment. Alcohol & Alcoholism (Suppl 1): 455-459.

64. Bitran M and Kalant H (1991). Learning factor in rapid tolerance to ethanol-induced motor impairment. Pharmacol. Biochem. Behav. 39: 917-922.

65. Lê AD, Poulos CX and Cappell H (1979). Conditioned tolerance to the hypothermic effect of ethyl alcohol. Science 206: 1109-1110.

66. LeBlanc AE, Matsunaga M and Kalant H (1976). Effects of frontal polar cortical ablation and cycloheximide on ethanol tolerance in rats. Pharmacol. Biochem. Behav. 4: 175-179.

67. Speisky MB and Kalant H (1987). Effects of anisomycin, PLG and DGAVP on rapid tolerance to ethanol in the rat. Xth Intl. Congr. Pharmacol. Sydney, Australia, Abstracts, P 781.

68. Van Wimersma Greidanus TjB, Bohus B and De Wied D (1975). The role of vasopressin in memory processes. In Gispen WH, van Wimersma Greidanus TjB, Bohus B and De Wied D (eds.), Hormones, Homeostasis and the Brain. Amsterdam: Elsevier. pp. 135-141.

69. Hoffman PL, Ritzmann RF, Walter R and Tabakoff B (1978). Arginine vasopressin maintains ethanol tolerance. Nature 276: 614-616.

70. Lê AD, Kalant H and Khanna JN (1982). Interaction between des-Gly^9-[Arg^8]-vasopressin and serotonin on ethanol tolerance. Eur. J. Pharmacol. 80: 337-345.

71. Popham RE (1956). The Jellinek alcoholism estimation formula and its application to Canadian data. Quart. J. Stud. Alcohol 17: 559-593.

72. Popham RE (1960). The applicability to Finnish mortality data of a trend method for the estimation of P values. In K Bruun, E Koura, RE Popham and JR Seeley (eds.): Liver cirrhosis mortality as a means to measure the prevalence of alcoholism. Helsinki: Finnish Foundation for Alcohol Studies. pp. 51-70.

73. Schmidt W (1977). The epidemiology of cirrhosis of the liver: a statistical analysis of mortality data with special reference to Canada. In Fisher MM and Rankin JG (eds.), Alcohol and the Liver. New York: Plenum Publishing. pp. 1-26.

74. Babor TF, Mendelson JH, Greenberg I and Kuehnle J (1978). Experimental analysis of the "happy hour": effect of purchase price on alcohol consumption. Psychopharmacology 58: 35-41.

75. Ledermann S (1964). Alcool, Alcoolisme, Alcoolisation: Mortalité, Morbidité, Accidents du Travail. Inst. Nat. d'Etudes Démog, Travaux Doc. Cahier 41. Paris: Presses Universitaires de France.

76. Smart RG, Whitehead P and Laforest L (1971). The prevention of drug abuse by young people: an argument based on the distribution of drug use. Bull. Narc. 23(2): 11-15.

77. Sullivan WT, III (1984). Karl Jansky and the discovery of extra-terrestrial radio waves. In Sullivan WT, III (ed.), The Early Years of Radio Astronomy. Cambridge: Cambridge University Press. pp. 3-42.

78. ScienceScope. Science: 255: 391, (1992)

79. Pavitt K (1991). What makes basic research economically useful? Research Policy 20: 109-119.

80. Franko LG (1989). Global corporate competition: Who's winning, who's losing, and the R & D factor as one reason why. Strategic Management J. 10: 449-474.

81. McMillan CJ (1991). Services: Japan's 21st Century Challenge. Ottawa: Canada-Japan Trade Council. Appendix A-13.

708225

BINDING SERVICES

NOV 1 7 1993

HV 5000 C3 W56 1992
WINDOWS ON SCIENCE 50440959

UTL AT DOWNSVIEW
D RANGE BAY SHLF POS ITEM C
39 20 19 21 12 006 8